INVESTIGACION y CIENCIA

ORQUIDEAS UNISEXUALES

AGUJEROS NEGROS Y FORMACION DE LAS GALAXIAS

POSIBILIDADES REALES DE LA TERAPIA GENETICA



NANOTECNOLOGIA

ENERO 1991
600 PTAS.



Enero de 1991 Número 172



Edición española de SCIENTIFIC AMERICAN

SECCIONES

COLABORADORES DE ESTE NUMERO

Asesoramiento y traducción:

José M. García de la Mora: *La ciencia y la técnica, tras el milagro de Occidente*; Ramón Pascual: *Agujeros negros en los centros galácticos*; Santiago Torres: *Terapia génica*; Ana Mª García: *El significado de los sueños*; Luis Bou: *Teoría de nudos y mecánica estadística*, *Juegos de ordenador* y *Juegos matemáticos*; Antonio Blanco: *Una fábrica romana*; Amando García: *Nanotécnica*; José Vilardell: *Hace...*

Ciencia y sociedad:

Josep-Enric Llebot y Joandoménec Ros

Ciencia y empresa:

Manuel Puigcerver

Libros:

A. Prevosti, J. Bartolomé, A. Domínguez, M. Durfort y L. Alonso

PROCEDENCIA DE LAS ILUSTRACIONES

Página	Fuente
5	Granger Collection
6	Photo Researchers, Inc.
7	© Mary Evans Picture Library, Photo Researchers, Inc.
8	Bettman Hulton
10	Granger Collection
11	General Electric Company
15	George Retseck
16	Rick Perley, Observatorio Nac. de Radioastronomía (*arriba*), George Retseck (*abajo*)
17	Observatorio Palomar, Inst. de Tecnología de California (*arriba*), Steve Kent (*centro*), George Retseck (*abajo*)
18-20	George Retseck
21	Fred Lo
25	Gamma-Liaison Network
26-28	Dana Burns-Pizer
29	James M. Wilson y J. Roy-Chowdhury
30	Dana Burns-Pizer
45	Marc Chagall, Scala/Art Resource
46	Carol Donner
47	Patricia J. Wynne
48	Carol Donner (*arriba*), Gabor Kiss (*abajo*)
49	Gabor Kiss
50	Carol Donner (*arriba*), Gabor Kiss (*abajo*)
53-57	Michael Goodman
61	A. Trevor Hodge
62-65	Tom Prentiss
67	Gustavo A. Romero
68	G. A. Romero y Pere Rovira
70	Gustavo A. Romero
71-75	G. A. Romero y Pere Rovira
76-77	Edward M. Clausen, Jr., y James P. Harbison, Bell Communications Research
78	John R. Deecken
79	Spectra Diode Laboratories, (*ambas fotografías*), John R. Deecken (*dibujo*)
80	Kim Y. Lee, IBM (*izquierda*), David M. Hwang, Bell Communications Research (*derecha*) y John R. Deecken
81	John McGrail
82	Henryk Temkin, AT&T Bell (*izquierda*), Kim Y. Lee, IBM (*derecha*), John R. Deecken
83	John N. Randall, Texas Instr. (*izq.*), Galen D. Stucky (*derecha*) y John R. Deecken
84	Edward Bell
86	Russell F. Smith
88-90	Edward Bell
92	VMBO
93-96	PLS



La microfotografía electrónica de transmisión que se reproduce en la portada muestra la estructura de un semiconductor de fosfuro de indio que contiene un punto cuántico. Construido por Henryk Temkin y sus colaboradores de los Laboratorios AT&T Bell, el tamaño de este elemento es menor que la milésima parte del grosor de un cabello humano en su base. Cuando se aplica un voltaje a esta estructura, los electrones resultan atrapados en la profunda región de color púrpura que se observa en el cuello del dispositivo (fabricado con arseniuro de galio e indio).

INVESTIGACION Y CIENCIA

DIRECTOR GENERAL Francisco Gracia Guillén
DIRECTOR EDITORIAL José María Valderas Gallardo
DIRECTORA DE ADMINISTRACIÓN Pilar Bronchal Garfella
PRODUCCIÓN César Redondo Zayas
SECRETARÍA Purificación Mayoral Martínez
EDITA Prensa Científica, S. A. Viladomat, 291 6º 1ª - 08029 Barcelona (ESPAÑA)
Teléfonos 321 81 91 - 321 83 48 Telefax 419 47 82

SUSCRIPCIONES

Prensa Científica S. A.
Viladomat, 291 6º 1ª
08029 Barcelona (España)
Teléfonos 321 81 91 - 321 83 48

Precios de suscripción, en pesetas:

	Un año	Dos años
España	6600	12.000
Extranjero	7300	13.400

Ejemplares sueltos:

Ordinario: 600 pesetas
Extraordinario: 775 pesetas

— Todos los precios indicados incluyen el IVA, cuando es aplicable.
— En Canarias, Ceuta y Melilla los precios incluyen el transporte aéreo.
— El precio de los ejemplares atrasados es el mismo que el de los actuales.

DISTRIBUCION

para España:

MIDESA
Carretera de Irún, km. 13,350
(Variante de Fuencarral)
28049 Madrid Tel. 652 42 00

para los restantes países:

Prensa Científica, S. A.
Viladomat, 291 6º 1ª - 08029 Barcelona
Teléfonos 321 81 91 - 321 83 48

PUBLICIDAD

Madrid: Gustavo Martínez Ovín
Menorca, 8, bajo, centro, izquierda.
28009 Madrid
Tel. 409 70 45 - Fax 409 70 46
Cataluña: Marcel Klein
M. K. Publicidad
Ortigosa, 14-16, 3º, D. 20
08003 Barcelona
Tel. 268 45 05 - Fax 268 16 07

OJD DIFUSION · PUBLICIDAD

ISSN 0210-136X Dep. legal: B. 38.999-76

Fotocomposición: Tecfa. Línea Fotocomposición, S.A. Almogàvers, 189 - 08018 Barcelona
Fotocromos reproducidos por T2, Llull, 145 - 08018 Barcelona
Imprime Rotographik, S.A. Ctra. de Caldes, km 3,7 - Santa Perpètua de la Mogoda (Barcelona)

Printed in Spain - Impreso en España

Hace...

...cincuenta años

Scientific American: "Los premios Nobel empezaron a concederse en 1901 y de ellos Alemania ha cosechado el mayor número. Ahora bien, mediante una estimación basada en la población relativa de cada país, y limitada a la última década, puede atribuirse a Suiza una puntuación de 4,65, seguida de Holanda con 2,17 y, tras ambas, Gran Bretaña con 1,75, Austria con 1,62, Alemania con 1,35, Estados Unidos con 0,78, Francia con 0,44 e Italia con 0,22. La puntuación estadounidense de 0,78, alcanzada durante los dos últimos lustros, *es superior a cuatro veces* la de 0,18 que se anotó en los primeros años."

"Los casos recalcitrantes de esquizofrenia, que es la forma más común de demencia, han reaccionado alentadoramente ante un nuevo tratamiento quirúrgico en el que se seccionan fibras del lóbulo frontal del cerebro. Esta operación recibe el nombre técnico de leucotomía prefrontal, aludiendo al corte de sustancia blanca de los lóbulos prefrontales del cerebro. Ideada por el doctor Egas Moniz, de España, fue introducida en EE.UU. por los doctores Walter Freeman y James W. Watts, de la facultad de medicina de la Universidad George Washington."

"Los cazadores y tiradores al plato 'adelantan la puntería', es decir, apuntan por delante de los blancos, pero disparan las armas desde una posición de inmovilidad casi total. Los artilleros aéreos, por su parte, deben cazar unas presas que vuelan a 400 millas por hora embarcados en unas aeronaves que se desplazan a velocidades del mismo orden. Un artillero que dispare desde un avión que esté derrapando, por ejemplo, deberá tener en cuenta que el arma que maneja se mueve *en tres direcciones a la vez*: hacia adelante, de costado y hacia abajo. Para ayudarse en la puntería, estos artilleros emplean unos visores que se balancean excéntricamente a fin de compensar el torbellino de la hélice. Hay, no obstante, otro fenómeno, llamado 'efecto Magnus', que ningún visor compensa. Se debe aquél al giro en sentido horario con que los proyectiles salen del cañón. Cuando un artillero rota su arma para lanzar una ráfaga por la izquierda, el giro interno de los proyectiles hace que éstos tiendan a descender en su trayectoria, del mismo modo que una pelota de tenis que haya recibido un efecto de sobregiro. Y al revés, en una ráfaga por la derecha, las balas tienden a subir en su trayectoria, como una pelota de tenis con efecto de contragiro."

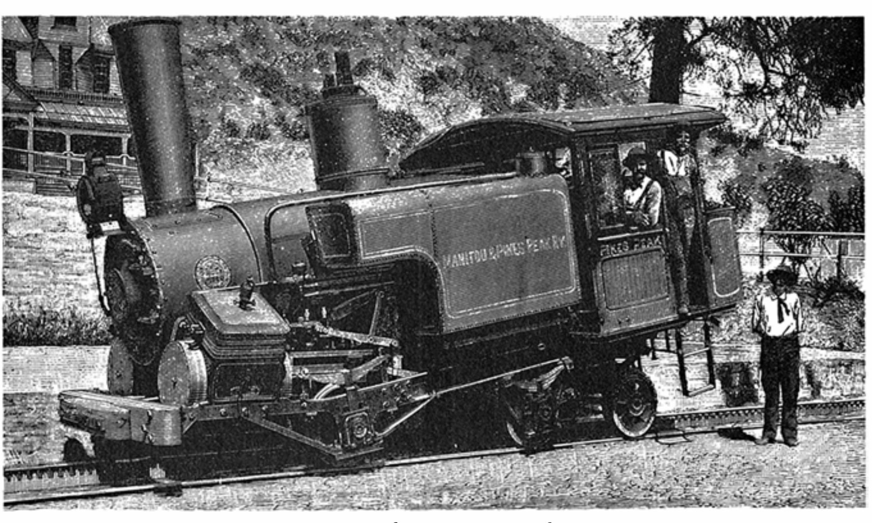
"El registro sonoro de la película *Fantasía*, de Walt Disney, no puede emplearse con los equipos actuales de las salas de exhibición. El margen dinámico de la mejor producción cinematográfica con sonido en película es de unos 35 decibelios. El de una orquesta sinfónica es del orden de los 70 decibelios. En el 'Fantasound', que es como se conoce al nuevo sistema de registro, el margen dinámico alcanza los 75 decibelios. En las películas convencionales, el registro sonoro lo lleva una estrecha banda longitudinal. Esa pista la tiene también 'Fantasía', pero sólo para casos de emergencia. La calidad procede de una segunda banda del tipo normal de película que pasa por una unidad reproductora de sonido especialmente diseñada. Esta película funciona en simultaneidad con la película de fotogramas y exactamente sincronizada con ella. La segunda, o película-de-sólo-sonido, lleva *cuatro* pistas de sonido separadas."

...cien años

Scientific American: "Hace varios años que los inventores pusieron a trabajar su ingenio para idear un aparato de dimensiones reducidas que permita a los aficionados tomar fotos sin que nadie se percate. Ya tenemos algo destinado a encontrar mucho éxito entre los practicantes de la fotografía. Se trata de una corbata con pasador. Este último es el objetivo y la primera es la cámara. Cuando alguien se nos aproxime y nos hable a una distancia de cincuenta centímetros, e incluso de un metro, apretamos una perilla que escondemos en el bolsillo y ya tenemos su retrato."

"El doctor Henry Schliemann nació en 1822 y tempranamente se sintió inclinado hacia el estudio del griego, afirmándose que cuando era apenas un niño tomó la decisión de descubrir Troya. Sus actividades comerciales prosperaron y acabó haciendo fortuna. Dedicó entonces el resto de su vida a la investigación arqueológica en Grecia y Asia Menor. En 1874 apareció su 'Troya y sus restos', en donde exponía los resultados de sus excavaciones en el emplazamiento de la antigua Troya. La vida familiar del doctor Schliemann discurrió bajo la influencia de sus gustos clásicos. Así, su esposa era griega y en su casa se hablaba en griego; hasta los criados recibían nombres helenos. A sus dos hijos se les impusieron los nombres de Agamenón y Andrómaco. En el curso de sus correrías por el mundo dio en hallarse en California cuando ésta fue admitida en la Unión y él mismo se hizo ciudadano estadounidense. El día 27 de diciembre fue anunciada su muerte en Nápoles."

"Estas máquinas de tren fueron fabricadas por los Talleres de Construcción de Locomotoras Baldwin, de Filadelfia, y presentan un aspecto muy diferente al de las locomotoras ordinarias. Cada una pesa 23,5 Tm y están construidas de modo que queden horizontales sobre una pendiente del 16 por ciento. La tracción la transmite una corona con un dentado similar al de los piñones de la vía."



Locomotora de montaña Pike

La ciencia y la técnica, tras el milagro de Occidente

La estrecha vinculación entre el auge del conocimiento científico y el avance técnico ha permitido a las economías libres de las naciones occidentales alcanzar una prosperidad sin precedentes

Nathan Rosenberg y L. E. Birdzell, Jr.

La desigualdad económica entre las naciones es hasta cierto punto una creación de los últimos 250 años. A mediados del siglo XVIII, el habitante medio de Europa occidental tenía un bienestar material no demasiado diferente del súbdito chino o, para el caso, del griego o romano de la antigüedad. Sólo una parte minoritaria de la población disfrutaba de una renta apreciablemente superior al mínimo requerido para el sustento vital. Las elites de una nación tenían pocas razones para envidiar a las de otra.

Mas, alrededor de 1800, fue haciéndose notorio que, si en Europa iba en aumento la minoría que contaba con ingresos superiores a los del nivel de mera subsistencia, ello se debía, al menos en parte, a que la ciencia y la técnica progresaban aquí más deprisa que las de cualquier otra zona del mundo. Al aumento del número de fábricas y a la expansión de la energía mecánica se dio en llamarlos revolución industrial. El proceso de desarrollo y de transformación se aceleró durante el siglo XIX y ha continuado a lo largo del XX. A este período sin igual de crecimiento económico prolongado, que ha hecho al mundo occidental manifiestamente más rico y poderoso que al resto, lo llaman a veces los historiadores "el milagro de Occidente".

NATHAN ROSENBERG y L. E. BIRDZELL, Jr. publicaron en 1986 una obra sobre el milagro de Occidente que llevaba por título *How the West Grew Rich* ("Así se enriqueció Occidente"). Rosenberg, experto en desarrollo económico y técnico, da clases en la Universidad de Stanford. Birdzell, de la fiscalía del Estado, enseña en Rhode Island.

Describir este fenómeno simplemente como desarrollo económico prolongado no proporciona una idea exacta de sus auténticas dimensiones. De mediados del siglo XVIII al día de hoy, la renta per cápita se ha decuplicado.

La población de Europa se ha quintuplicado y la de los los Estados Unidos se ha multiplicado por 80. Disminuyó drásticamente la mortalidad infantil y se ha doblado la esperanza de vida. Se eliminó el hambre y desaparecieron las pestes. La producción de alimentos, que en algunos países había ocupado nada menos que al 90 por ciento de la población activa, no llegaría a ocupar finalmente ni al 5 por ciento. En el siglo XIX la urbanización avanzó al compás de las nuevas técnicas: instalaciones sanitarias, construcción, comunicaciones, distribución de energía y tantos otros servicios. La urbanización y el aumento de los ingresos introdujeron cambios en el nivel de vida y de salud, en el trabajo, en los valores y en otros aspectos de la existencia personal, familiar y comunitaria.

Cabría suponer que los historiadores hubieran aclarado desde hace ya mucho las razones por las que se produjo el milagro de Occidente, pero lo cierto es que el fenómeno no ha sido examinado con todo el detenimiento que se merece. Por eso, cuando hace 40 años los países menos desarrollados del Tercer Mundo volvieron sus miradas hacia Occidente en demanda de ayuda para elevar sus rentas per cápita, muchos de los asesoramientos que recibieron reflejaban una inadecuada comprensión de la manera en que Occidente había alcanzado la prosperidad de que disfruta. Muy recientemente, la indudable determinación de la Unión Soviética y de los países del Este, resueltos a acabar con la abismal diferencia entre su renta y la de Occidente, ha vuelto a poner en el candelero esta cuestión: ¿cuáles son las causas del desarrollo occidental con las que no contaron los países menos desarrollados y los que siguieron la doctrina socialista?

Se han propuesto diversas explicaciones populares del milagro de Occidente. No faltan quienes lo atribuyen al imperialismo, aunque muchos de los países de mayor éxito económico eran ya prósperos antes de recurrir al imperialismo y algunos países ricos, como es el caso de Noruega y Suiza, nunca adoptaron políticas expansionistas. Se da también el fenómeno contrario: varias de las más formidables potencias imperialistas, como fueron España y Portugal, decayeron rápidamente y sus economías quedaron anémicas.

Otras teorías relacionan la riqueza con la posesión de recursos naturales. Pero esos recursos no tienen relevancia económica mientras no se disponga de los conocimientos y los medios precisos para utilizarlos. Los pueblos precolombinos de América del Norte tenían más o menos los mismos recursos naturales que los habitantes de hoy. Japón, que posee muchos menos recursos naturales que Indonesia, México o la Unión Soviética, ha logrado un éxito mucho mayor en su crecimiento económico. La historia moderna de ciudades-estado (citemos Hong-Kong y Singapur, por no mencionar Venecia, que sólo contaba como recursos naturales con una laguna pantanosa) contradice toda explicación que sólo contemple los recursos naturales.

Lo malo del caso es que estas hipótesis han influido en la política gubernamental de los países en vías de desarrollo. Centrando su atención en

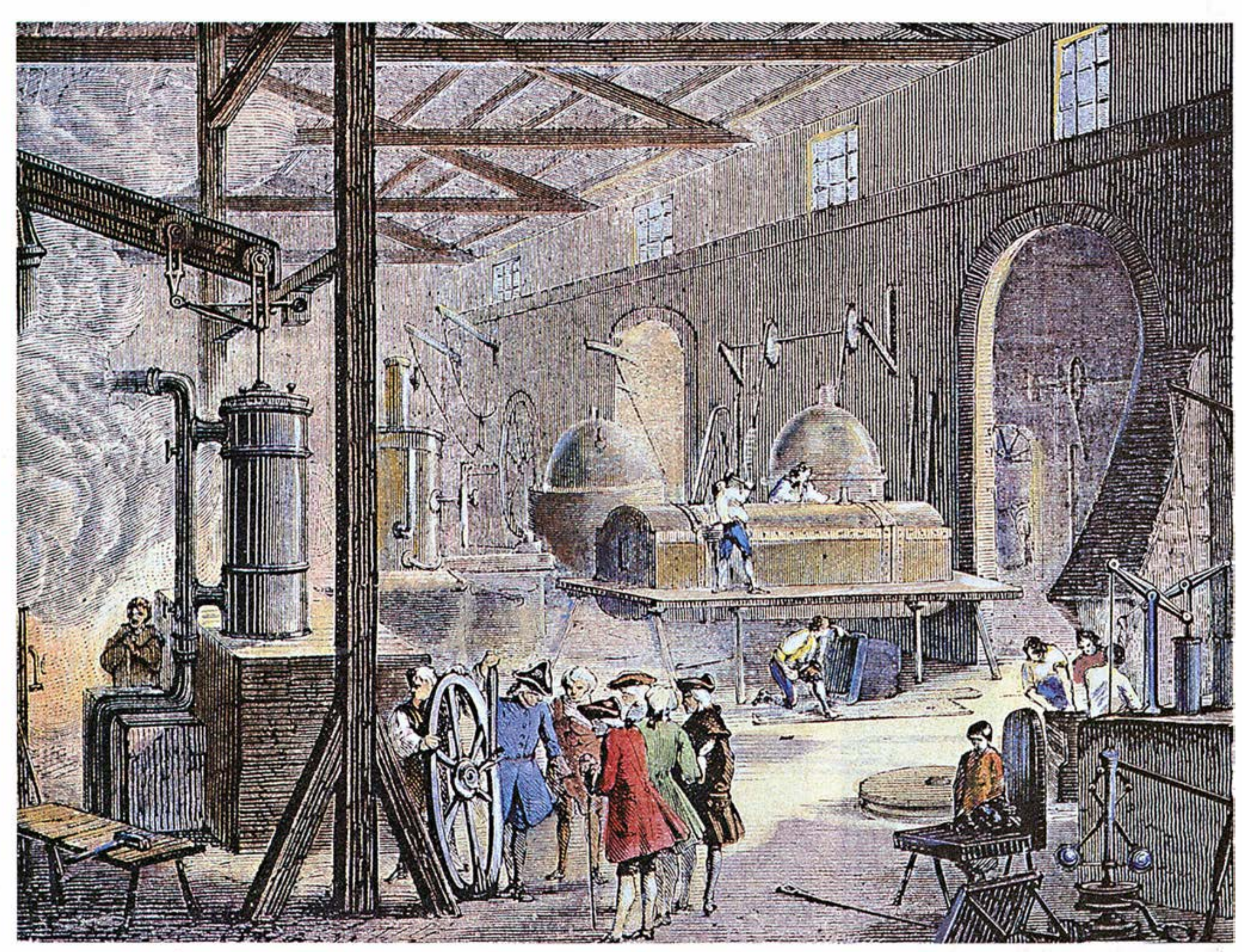
factores de discutible importancia, han descuidado cambios institucionales que podrían haber abierto cauces al desarrollo, en especial los que hubieran dado acceso a técnicas superiores.

La tecnología occidental se desarrolló primordialmente en la esfera económica, y ha sido considerada a menudo mero resultado de las necesidades e instituciones económicas. La ciencia, en contraste, tuvo unos orígenes más complejos y difícilmente se podría limitarla a pura respuesta automática a las condiciones económicas. Durante mucho tiempo, la ciencia contribuyó poco al desarrollo económico y a la técnica industrial. Cuando a mediados del siglo XIX escribía Marx sus obras, las "colosales fuerzas de producción" que veía funcionando habían sido creadas primordialmente por gentes que se ocupaban en la industria, con escasa contribución de quienes hoy llamaríamos científicos. La habilidad mecánica y el ingenio del que nació la maquinaria de precisión y todo el instrumental de laboratorios, talleres y factorías de los siglos XVIII y XIX provenían más del arte de los relojeros y de los pulidores de lentes que de la ciencia.

Ahora bien, desde 1880, más o menos, la técnica industrial se ha alimentado mucho más de fuentes científicas ajenas a la industria. Con el éxito de los esfuerzos por hacer encajar los fenómenos naturales en estructuras teóricas inaccesibles sin una formación especializada, los ingenieros industriales que la poseen se han convertido en transmisores y usuarios del conocimiento y los métodos científicos. Y, lo que es más, a lo largo de los últimos cien años la industria creó laboratorios de investigación capaces de ampliar las estructuras teóricas de la ciencia. Aunque ésta nació como institución ajena a la esfera económica, a lo largo del siglo XX su avance se ha hecho inseparable del de la técnica y la economía de Occidente.

Para explicar el milagro económico de Occidente y su relación con la ciencia, hemos de considerar ante todo algunas de las causas del gran éxito alcanzado por la ciencia occidental, logro que merece por sí solo el título de "milagro". Una de tales causas es que la ciencia occidental se ha organizado mejor para arrancarle sus secretos a la naturaleza y ha empleado en ello mayores recursos que los utilizados por la ciencia en otras culturas.

Mucho tiempo después de que, a finales del siglo XV, se divulgara la imprenta, la investigación científica seguía siendo principalmente una actividad descentralizada, individual incluso, en la que los científicos se comunicaban ocasionalmente en letra impresa o en manuscritos sus descubrimientos. Los comienzos de la ciencia occidental no fueron, desde luego, un fenómeno local: su cuna se extiende desde la Polonia de Copérnico has-



1. LA FABRICA SOHO, en los aledaños de Birmingham, producía, en el siglo XVIII, piezas para máquinas de vapor. Los socios de la empresa, James Watt (padre de las máquinas de vapor eficientes) y Matthew Boulton, eran miembros de la Sociedad Lunar de Birmingham, agrupación formada por hombres de negocios, inventores y científicos. Por todo Occidente florecieron en aquella época parecidas asociaciones de científicos y empresarios que fomentaban simultáneamente el progreso científico, el técnico y el económico. (Grabado publicado por cortesía de Granger Collection.)

ta la Dinamarca de Tycho Brahe, desde la Italia septentrional de Galileo hasta la Bohemia de Kepler, la Francia de Descartes y Lavoisier, y la Inglaterra de Boyle y Newton.

Los primeros logros de la ciencia occidental se concentraron en la astronomía. La aparición de una comunidad científica europea relevante cuyos fines fueran más allá de la astronomía, no tuvo lugar hasta el siglo XVII. En 1660 se fundó la "Real Sociedad de Londres para el Fomento del Conocimiento Natural" (la "Royal Society"), con el fin de intercambiar opiniones sobre los informes de los muchos individuos que a la sazón estaban ocupados en investigaciones científicas. En el siglo XVII y en el XVIII se formaron otras muchas sociedades parecidas, creándose así una red de relaciones entre los científicos de Europa, que intercambiaban información no sólo entre sí, sino también con un lejano norteamericano llamado Benjamin Franklin, cuyos experimentos habían demostrado que el rayo era una descarga eléctrica.

Estas sociedades y sus publicaciones periódicas difundían la investigación más reciente y la pasaban por el tamiz de la crítica para someterla a los cánones científicos. Sus puntos de vista indicaban el camino a seguir en cada momento y servían de referencia para nuevas investigaciones que pudieran obtener el reconocimiento y el aplauso de otros científicos. Lo que ciertamente no ofrecían eran medios de ganarse la vida. Así, en 1695, Isaac Newton hubo de ver limitado su ascenso académico en Cambridge porque no había recibido las órdenes sagradas; para recompensarle por sus contribuciones científicas con algún medio de subsistencia, el gobierno británico tuvo que concederle un cargo fuera de la comunidad científica: se le nombró guardián de la Casa de la Moneda.

Aunque la idea de reunir a los científicos para que investigaran organizadamente en un centro equipado con instrumental de laboratorio y una adecuada biblioteca ya se había puesto en práctica en la primera mitad del siglo XV, bajo los auspicios de don Enrique el Navegante de Portugal, sólo a comienzos del siglo XIX se hizo habitual. En Londres, Sir Joseph Banks, el conde de Rumford y algunos otros miembros de la Regia Sociedad fundaron en 1799 la "Royal Institution" (Regia Institución) para que sirviera de laboratorio donde los científicos pudiesen trabajar en equipo y dar conferencias. Michael Faraday, un siglo posterior a Newton, labró toda su carrera en la Regia Institución; allí descubrió la inducción electromagnética.

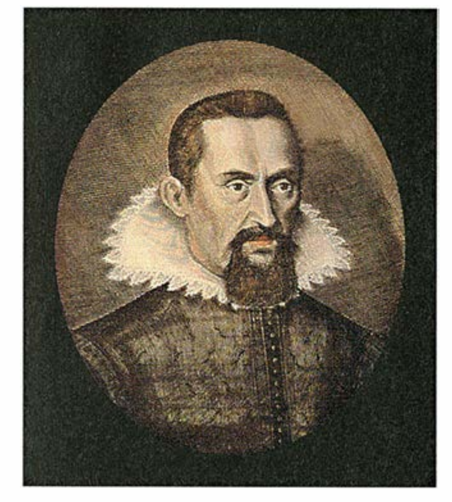
Por doquier fueron surgiendo instituciones parecidas. En 1795 se fundó en Francia la "École Polytechnique" ("Escuela Politécnica"). En los Estados Unidos la Universidad de Yale estableció la sección de ciencias de Sheffield en 1847 y el Instituto de Tecnología de Massachusetts abrió sus puertas en 1865. La ciencia fue así desarrollando gradualmente su propia investigación, y los estudiosos más afamados pudieron ser recompensados con nombramientos para ocupar cátedras y con diversos galardones académicos.

En los albores del siglo XIX, la ciencia occidental se había fragmentado en departamentos especializados: matemáticas, astronomía, física, química, geología, botánica, zoología y los estudios médicos de anatomía y fisiología. Algunos de ellos, como el de física, se subdividían aún en otros más específicos.

La ciencia occidental se había convertido en una institución que contaba con una amplísima finalidad general (la de explicar los fenómenos naturales), una división del trabajo en departamentos especializados con sus propias finalidades subsidiarias, una red de información que tenía a sus miembros al tanto de todos los progresos, un sólido sistema de crítica y revisión para evaluar los trabajos nuevos y dirimir los conflictos, centros dedicados en toda regla a la docencia y a la investigación, y una serie de premios para recompensar al trabajo que fuese estimado favorablemente por la comunidad científica.

Factor fundamental que mantenía firme y unida a la empresa era el de la adopción de un único patrón de verdad científica basado en la observación, el razonamiento, el experimento y su posibilidad de reproducción. Este patrón permitía a los científicos hacer uso de los hallazgos logrados en otros laboratorios, incluso de los de otras disciplinas. Y permitía también que artesanos, comerciantes, fabricantes y el resto de la población laboral aplicara los des-









2. ILUSTRES CIENTIFICOS contribuyeron en toda Europa al crecimiento del saber básico que estimuló el progreso técnico industrial y el desarrollo económico. Se destacaron entre ellos *(de arriba abajo y de izquierda a derecha)* Nicolás Copérnico, Juan Kepler, Sir Isaac Newton y Galileo Galilei.



3. SENADORES VENECIANOS escrutando los satélites de Júpiter a través de un telescopio bajo la dirección de Galileo. La astronomía fue uno de los primeros aspectos en los que centró su interés la ciencia, pero antes de que acabase el siglo XVII Europa empezaba a contar con una comunidad amante de la física, la medicina y otras ramas del conocer. Poco a poco, sin embargo, la ciencia fue dejando de ser tarea de individuos solitarios.

cubrimientos científicos a su ámbito profesional.

Su organización y sus amplias miras no son, claro está, las únicas causas del florecimiento de la ciencia occidental. Recibió también en herencia el vasto legado intelectual de anteriores civilizaciones: un alfabeto fonético, un sistema de cálculo árabe (o posiblemente indio) que incluía el cero como número, unas matemáticas que abarcaban geometría y álgebra, y una religión que había liberado a la naturaleza del animismo. Pero no fue Occidente un heredero perezoso, pues a comienzos del siglo XVIII había añadido ya unos cuantos monumentos intelectuales de su propia cosecha.

La del cálculo infinitesimal, por ejemplo, fue a todas luces una contribución decisiva. Otra, y puede que la de mayor peso, fue el desarrollo del método científico, que desde la época de Galileo incluyó una refinada forma de experimentación sistemática. Los hombres de ciencia e inventores helenísticos, islámicos y chinos sabían valerse de experimentos para poner a prueba y confirmar sus ideas, pero no parece que lograsen nada como los experimentos de Galileo con el plano inclinado, en los que las condiciones se iban variando sistemáticamente para averiguar cómo obraba la naturaleza. Ni tampoco se anticiparon a los inteligentes experimentos de Newton, en los que para explicar los fenómenos reales se ponen en juego fenómenos ideales (como el movimiento en el vacío).

Sin experimentación sistemática, el progreso científico y técnico procede lento y vacilante. Tomemos el caso del arado; revestía capital importancia en las sociedades predominantemente agrícolas, pero los inventos que permitían mejorar este apero se introducían con intervalos de siglos. Antes de la edad de la ciencia, no parece que nadie tratara de mejorar la forma del arado comparando la eficacia con que los distintos tipos de reja pudieran hendir las diferentes texturas del suelo. Otro ejemplo: los romanos ya conocían el cemento, pero se usó poco como material de construcción hasta que, a finales del siglo XIX, los químicos investigaron experimentalmente, variando de modo sistemático la mezcla de sus ingredientes, la manera en que resultara más aprovechable en sus aplicaciones estructurales. En pocas décadas, el cemento, incluido el hormigón armado, se convirtió en el material de construcción más empleado (en peso) del mundo occidental.

Quizás el extremo de mayor importancia respecto a la ciencia y la técnica occidentales estribe en su plena imbricación. En otras civilizaciones, las técnicas económicamente útiles dependían poquísimo, si es que dependían algo, de los saberes de astrónomos (o astrólogos), filósofos, matemáticos y demás sabios. Estos apenas si tenían algo que ofrecer a los labradores, a los marinos, a los herreros y a otros artesanos que habían desarrollado sus técnicas siguiendo sus tradiciones gremiales.

De hecho, los pensadores se encerraban a menudo ellos mismos en la torre de marfil de un mundo de ideas abstractas, escapándose del transitorio e imperfecto mundo de la realidad. En cambio, para los científicos occidentales no había vía de escape: sus métodos experimentales requerían meterse de lleno en el mundo real. Precisamente por haberse ocupado tanto de la realidad es por lo que los científicos lograron realizar tantas cosas.

Los países del Tercer Mundo y los

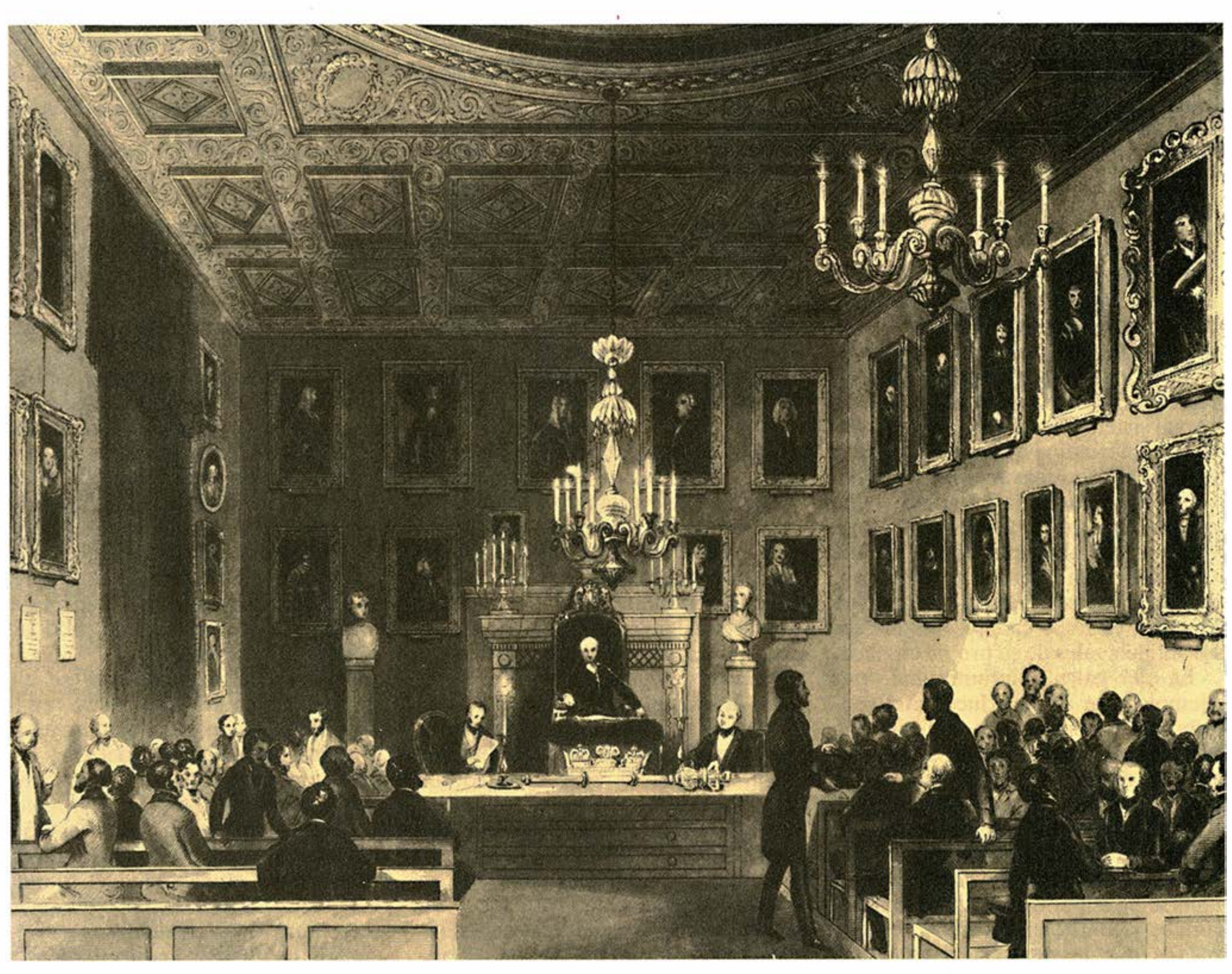
del Este europeo van económicamente muy por detrás de los de Occidente, a pesar de que tienen acceso al patrimonio mundial de los conocimientos científicos. Y, por cierto, la Unión Soviética ha contribuido sustancialmente al enriquecimiento de tal patrimonio. Lo que a esos países pobres parece faltarles es la capacidad de Occidente para traducir el conocimiento científico en productividad económica, capacidad que depende también de las características de los individuos y de las instituciones de cada nación. La técnica, que es la intermediaria entre el conocimiento científico y la actividad económica, se desarrolla a partir de las necesidades y de las instituciones locales; su aplicación económica correcta supone algo más que un sistema de enseñanza.

Para que haya crecimiento económico se requiere, sin duda, el acierto en adaptar las técnicas productivas a las necesidades locales. Sean cuales fueren los orígenes de una técnica determinada, los individuos y las instituciones que la empleen deberán ser capaces de entenderla y de experimentar con ella, así como de valorar las repercusiones económicas que derivan de su utilización.

Japón, al que se suele considerar el primer país equiparable en sus logros económicos a Occidente, nos brinda un instructivo ejemplo de instituciones que aplican las técnicas como es debido. La modernización nipona, que comenzó a finales de los años sesenta del siglo pasado, procuró ante todo, en sus primeras fases, la mejora de la agricultura. Inicialmente las técnicas, tomadas en préstamo de Norteamérica, consistieron en intensificar el capital y buscaban maximizar la productividad por trabajador. Pero los japoneses no tardaron en darse cuenta de que las técnicas apropiadas para los Estados Unidos, donde tanto abundan las tierras cultivables, no lo eran para una economía como la de su país, con tan numerosa mano de obra y tan grave escasez de terrenos aptos para el labrantío. Consecuentemente cambiaron las técnicas, adoptando esta vez las de Europa occidental, que consistían en una mayor intensificación del trabajo y maximizaban la productividad del más escaso recurso japonés: el suelo.

También en el sector secundario supo el Japón adaptar la técnica de Occidente a su economía rica en fuerza de trabajo y pobre en capital. Las fábricas japonesas adquirían frecuentemente maquinaria de segunda mano en vez de otra nueva. Siempre que era posible sustituían el capital por el trabajo y empleaban técnicas extranjeras adaptándolas para aumentar el peso específico de la mano de obra. Las hilaturas, por ejemplo, añadían turnos de trabajo extraordinarios y gastaban más mano de obra en mantener y reparar la maquinaria para prolongar su vida útil. Al tender una vía férrea, el número de obreros empleados por kilómetro de trazado en Japón era dos veces y media superior a los que hacían falta en Norteamérica.

Frecuentemente se pasan por alto los paralelismos que, en cuanto al de-



4. REGIA SOCIEDAD DE LONDRES para el Progreso y Fomento de la Ciencia Natural, fundada en 1660. Fue una de las primeras asociaciones que se formaron con el fin de favorecer la comunicación científica. Numerosas sociedades evaluaron los descubrimientos y los divulgaron.

sarrollo de la industria, se dan entre el Japón y los Estados Unidos, país éste donde ese desarrollo empezó varios decenios antes que en aquél. Como en Japón, los Estados Unidos adoptaron en un principio técnicas foráneas, sobre todo inglesas. La industrialización comenzó en Nueva Inglaterra, Nueva York, Pennsylvania y Delaware, regiones en las que, como en Japón, eran ya apreciadas la alfabetización y la educación básica. Los Estados Unidos tuvieron también que adaptar las técnicas importadas a las especiales circunstancias del país, lo mismo que ocurrió en el Japón.

Sin embargo, como los Estados Unidos poseían abundantes recursos naturales, sus adaptaciones supusieron que las técnicas industriales intensificaran más el empleo de recursos y menos el del trabajo. Los europeos que visitaban los Estados Unidos a mediados del siglo XIX criticaban a menudo a los americanos por su derroche de recursos naturales. Las técnicas agrícolas estadounidenses provocaban con frecuencia una rápida disminución de la fertilidad del suelo, mas como tierra no faltaba, las pérdidas podían soportarse. Los norteamericanos inventaron una maquinaria para trabajar la madera que a los británicos les pareció extremadamente despilfarradora: pero por entonces no carecía de buen sentido económico en un país tan rico en bosques.

Bien entrado ya el siglo XIX, diversas empresas privadas de Alemania y los Estados Unidos crearon laboratorios de investigación industrial con el fin de desarrollar nuevos productos y métodos de producción. En 1856 el químico inglés William Henry Perkin descubrió la malveína, primer colorante de anilina derivado del alquitrán de hulla que revestiría interés industrial, y al año siguiente instaló una factoría para su producción. Este trabajo señaló el comienzo de una importantísima rama de la química y, a la vez, el de una no menos floreciente industria. Perkin siguió adelante con otras notables aportaciones, pero un individuo aislado no podía abarcar todas las posibilidades que brindaba la química de los derivados del alquitrán. Para lograr tal cosa era preciso institucionalizar la capacidad inventiva. Sólo cuando los fabricantes alemanes de productos químicos fundaron unos cuantos laboratorios científicos en los que se podían hacer investigaciones organizadas y sistemáticas, surgió, entre los años 1890 y 1914, una importante industria en torno al alquitrán de hulla.

Los éxitos conseguidos por los laboratorios alemanes inspiraron, en los Estados Unidos, la fundación del laboratorio científico de la Compañía General Eléctrica. Al constituirse en 1892 la General Eléctrica, recayó al principio en Charles Steinmetz, refugiado político de talento procedente de Alemania, el cargo de inventor residente. Pero en 1900, tras haber producido los laboratorios germanos materiales de superior calidad para la fabricación de los filamentos de las bombillas eléctricas, la General Eléctrica le encargó a Willis R. Whitney, profesor de química en el Instituto de Tecnología de Massachusetts, que organizara un laboratorio en toda regla. Durante la primera mitad del siglo XX se multiplicó el número de laboratorios científicos filiales de empresas privadas y se crearon nuevas e importantes instituciones que contribuyeron bastante al desarrollo de los conocimientos básicos de la ciencia y de la técnica.

El mejor camino para el progreso tecnológico está a menudo mal señalizado. La gran virtud de las empresas privadas en las economías de mercado consiste en que se erigen en fuentes de decisiones independientes cuando se trata de explorar las fronteras de la técnica. Un individuo o una institución aislada no tienen autoridad para vetar unilateralmente una labor investigadora. La importancia de repartir entre muchos la toma de decisiones la ilustra el éxito que ha alcanzado el ordenador personal, del cual se dice que fue rechazado por el principal fabricante de ordenadores de los Estados Unidos. De modo parecido, la exportación de automóviles japoneses a Norteamérica la inició una compañía nipona en contra de lo que aconsejaba el gobierno de su país. A buen seguro que al gobierno no le faltaba razón, pues los primeros coches exportados habían sido hechos al uso japonés y sus prestaciones quedaban deslucidas en manos de los primeros clientes estadounidenses. No obstante, el exportador aprendió en seguida lo que requería el mercado de los Estados Unidos y tomó las necesarias medidas correctoras. Su ejemplo fue seguido provechosamente por otras compañías.

Dadas las incertidumbres científicas, técnicas y comerciales que toda innovación comporta, una economía que quiera ser eficiente ha de hallar un equilibrio entre el emprender sólo el más seguro de los proyectos y el lanzarse a realizar cada idea que llame a la puerta, y que demasiado a menudo sigue revoloteando por la imaginación cuando ya se ha desvanecido toda esperanza realista de lograr lo deseado. Con sus recompensas y penalizaciones, las economías de mercado han logrado, de manera muy satisfactoria hasta el presente, mantener el necesario equilibrio y transformar los saberes científicos y técnicos en bienes y servicios útiles.

A pesar de todo, la difuminación de autoridad que conlleva la economía para acometer proyectos innovadores no acabaría de funcionar si las esperanzas de obtener elevadas compensaciones no fueran refrenadas en quienes han de tomar la decisión por el riesgo de sufrir graves pérdidas. La atención pública suele fijarse en las recompensas, a veces sensacionales, de la innovación, pero el inversor prudente nunca olvida cuál es el papel que desempeñan en la innovación los tribunales que declaran las quiebras: entierran los fracasos.

En Occidente, si bien las aportaciones innovadoras se originan a menudo en firmas de rancio abolengo, muchas de las más importantes invenciones han sido comercializadas por nuevas empresas nacidas con ese fin o por otras ya antiguas pero completamente remodeladas. El papel que en lo tocante a la innovación desempeñan las nuevas empresas tiene su importancia, no sólo por lo que directamente aportan, sino por su implícita amenaza para las firmas arraigadas, que de otro modo podrían caer en la apatía y evitarse los riesgos de la innovación. Ahora bien, la libertad de empresa está muy limitada en los países socialistas y gravemente entorpecida en algunos en vías de desarrollo, donde para poner en marcha cualquier iniciativa hay que recabar el permiso de toda una legión de burócratas gubernamentales.

Además de la investigación y el desarrollo empresarial, los países de Occidente han ido cultivando durante todo el siglo XX otras fuentes de conocimiento económicamente útil, sobre todo en áreas en las que los mercados no han proporcionado adecuados incentivos. Los fondos públicos han llegado a ser decisivos para financiar la investigación cuando ésta se hace prohibitivamente costosa, como es el caso de la física de partículas. Los gobiernos corren también con los costes de la investigación en salud pública, medicina preventiva, tratamiento de enfermedades poco frecuentes y seguridad, campos éstos en los que es universal el deseo de alcanzar los objetivos, pero poco beneficio ofrecen a la iniciativa privada. Las investigaciones patrocinadas por los gobiernos se llevan a cabo tanto en centros públicos como privados. Desde la segunda guerra mundial es cada vez mayor el número de universidades

privadas que se han beneficiado del erario público, convirtiéndose en centros de activa investigación.

El desarrollo económico occidental se ha caracterizado por un aumento del comercio y del volumen de los mercados. El incremento del tráfico comercial estuvo vinculado en parte a los progresos técnicos de la navegación y a la aparición del ferrocarril, innovaciones que rebajaron los costos y disminuyeron los riesgos del transporte. A medida que se fue ampliando el ámbito del comercio de las especialidades regionales, los fabricantes alcanzaron mercados cuyo tamaño merecía la inversión pertinente en técnicas de producción en masa. Gracias a la instalación de cámaras refrigeradas en los barcos, por ejemplo, la creciente población europea pudo intercambiar sus productos manufacturados con carne de Argentina, Australia y Estados Unidos. Más recientemente, el volumen cada vez mayor de los mercados internacionales ha hecho que los fabricantes puedan permitirse suministrar una mayor variedad de productos, mejor adaptados a las necesidades particulares de los diversos países o de grupos de clientes. Esta tendencia se puede observar en mercados muy dispares: automóvil, prendas de vestir, comidas preparadas o aparatos electrónicos.

En los últimos años la importancia del comercio internacional ha quedado resaltada por las diferencias entre los avances en el desarrollo de los países que han competido de manera activa en el tráfico internacional de mercaderías y la situación de los que han adoptado políticas sustitutivas de la importación (proteccionismo) o de estrecha regulación del comercio, como es el caso de los países socialistas. No hay una explicación en la que todos coincidan al tratar de aclarar por qué la participación activa en el comercio internacional va tan íntimamente asociada con el crecimiento económico. Las posibles respuestas barajan la combinación de economías en equilibrio, mayores incentivos a la competitividad, economías de especialización y el que los gobiernos de los países exportadores se abstengan de imponer contraproducentes restricciones.

El sector fabril occidental nos ofrece un ejemplo muy interesante de especialización: el del creciente número de fabricantes que sólo producen determinados componentes o piezas, de manera señalada para las fábricas de automóviles y para las de aparatos eléctricos o electrónicos. Muchas marcas se limitan al diseño o a la venta del producto; a veces también al montaje final de los componentes que han fabricado los subcontratistas. Los avances técnicos concebidos por ingenieros americanos, alemanes o japoneses pueden, de este modo, crear empleo mediante la fabricación subcontratada no sólo en el propio país sino también en México, Corea del Sur, Taiwán, Singapur o en cualquier otro sitio. Las modernas redes de subcontratación son también redes de intercambio internacional, cuyas operaciones dependen por completo de unos medios de transporte, de comunicación y de procesamiento de datos que resultaban inimaginables en el año 1800.

Como en las economías de libre mercado las redes comerciales carecen de una autoridad central, su poder de organizar eficazmente actividades pasó inadvertido hasta mucho después de haberse iniciado el milagro de Occidente. Con su falta de autoridad central, la organización notoriamente ventajosa de las economías occidentales mantuvo el paso de la ciencia occidental, que, careciendo también de una autoridad suprema, llegó a ser no obstante una empresa eficiente, que coordinaba el trabajo de millares de científicos especializados y un amplio elenco de instituciones investigadoras. La descentralización reviste una importancia crucial para el crecimiento económico, pero además las sociedades de Occidente han gozado de un grado comparable



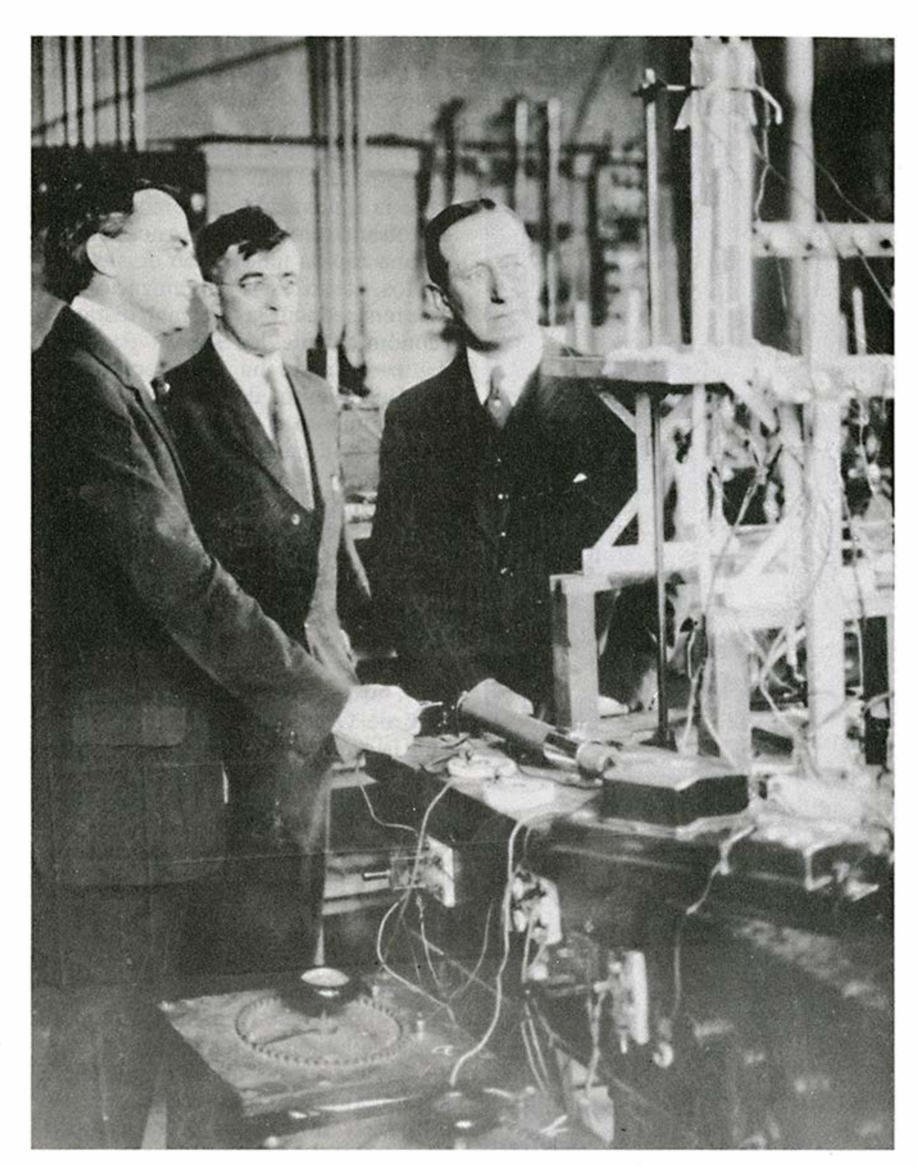
5. FERROCARRIL TRANSCONTINENTAL norteamericano. Terminó su tendido en Promontory Point (Utah) en 1869. Los progresos en la técnica del transporte impulsaron el desarrollo económico abriendo y ampliando mercados que, a su vez, permitían la mayor especialización de las fábricas.

de autonomía respecto al control político en otros frentes: el arte, la religión, la música, la literatura y demás importantes esferas sociales.

El problema básico del control de cualquier sistema económico radica en sentar las bases y fomentar un cambiante flujo de decisiones interdependientes que afectan a la producción y el consumo para elevar al máximo el bienestar humano. No mucho después del siglo XII empezó a insinuarse en las economías de la Europa occidental, tradicionalmente controladas por los gobiernos, las corporaciones gremiales y la Iglesia, un mercado de libre comercio. Comenzó un largo proceso, de siglos de duración, que comprendía la creación de nuevas ramas del comercio (entre ellas la del comercio internacional) fuera de la jurisdicción de las autoridades, y por otra parte se buscaba la liberación del control que éstas ejercían. Era tradicional en Occidente que las autoridades regularan manifiestamente los precios y los salarios, dentro de unos márgenes justos y equitativos, de acuerdo con criterios inevitablemente subjetivos.

El lento auge del intercambio mercantil libre de regulaciones fue reorganizando poco a poco aquellas economías mediante la transformación de los precios y los salarios, que pasaron a ser medios neutros desde el punto de vista ético, para mantener sincronizadas la oferta y la demanda, con unas consecuencias para la organización y el desarrollo de las economías occidentales que no acabarían de apreciarse hasta que las estudió Adam Smith a finales del siglo XVIII. Los organismos más tangibles de aquellos tiempos eran los gobiernos, los ejércitos y la Iglesia, todos ellos instituciones jerárquicas. Para muchos observadores, la idea misma de organización implicaba una jerarquía de autoridades, y su ausencia era equiparada al caos. No eran capaces de entender que, mediante unos mecanismos descentralizados, las economías occidentales estaban logrando implantar nuevos patrones de especialización y una organización cuyos alcances y cuya eficacia no tenían precedentes históricos.

Cuando el gran sociólogo alemán Max Weber escribió su penetrante obra sobre la organización, a comienzos del siglo XX, el poder organizador de los mercados era ya ampliamente reconocido entre los teóricos de la sociedad. Sin embargo, algunos consideraban a la burocracia una alternativa enteramente viable, que parecía ofrecer un posible retorno al viejo mundo de los precios y los salarios "justos", que a los socialistas y a otros muchos les resultaba atractivo. La comparación de lo que en los últimos setenta años han conseguido las economías de Occidente y las de los países socialistas, culminando éstas en las recientes agitaciones que han tenido lugar en la Unión Soviética y en la Europa del Este, sugiere que el juicio que las equiparaba dejaba de lado algo importante.



6. WILLIS R. WHITNEY, primer director del laboratorio de investigación de la empresa General Eléctrica *(izquierda)*, e Irving Langmuir, químico industrial galardonado con el premio Nobel *(centro)*, enseñan un aparato de producción de alto vacío a Guglielmo Marconi, el inventor de la radio *(derecha)*. Los nuevos laboratorios de las empresas ligaron estrechamente la técnica y la ciencia.

El tráfico mercantil puede, de suyo, ser dirigido por agentes gubernamentales, por gremios o sindicatos, consumidores, inversores o por otros grupos o individuos cualesquiera. Pero lo cierto es que en la historia de Occidente la introducción del mercado libre se debió a los esfuerzos de comerciantes que trabajaban en el marco funcional de una institución típica de Occidente: la empresa mercantil y manufacturera. Lo que parece más genuinamente occidental de estas instituciones es que con frecuencia han competido entre sí en términos de innovaciones de los productos, de su manufactura y de su distribución. El éxito de las economías occidentales en la asimilación de la técnica de Occidente no es consecuencia sólo del libre mercado, sino de que en el mercado se mueven firmas productoras que pueden ganar mucho si comercializan ideas nuevas antes que sus rivales.

Según hemos observado ya, la innovación es un modo arriesgado de competir, y las empresas que viven o mueren por las inseguras irregularidades de los beneficios y pérdidas que la innovación comporta se diferencian enormemente de los órganos admi-

nistrativos de cualquier gobierno. Este extremo es hoy una cuestión candente en la Unión Soviética y los países del Este, que, habiendo declarado ilegales los tipos de entidades mediante las que Occidente encauzaba la innovación comercial, se hallan ahora tremendamente necesitados de una rápida innovación de sus productos, fábricas y organización, si quieren reducir la abismal distancia económica que los separa de Occidente. Aunque ha ido creciendo el convencimiento de que es preciso confiar más en el libre mercado, hasta ahora no se ha hecho hincapié en la apremiante necesidad que hay de dejar que las empresas soviéticas y las todavía socialistas introduzcan innovaciones en sus productos, en sus métodos de producción y en su forma de organizarse y de que cosechen los beneficios y pérdidas resultantes.

Al buscar explicaciones del milagro de Occidente, hemos sostenido que la expansión económica a largo plazo y la expansión técnica van a la par, puesto que ninguna de ellas se ha dado durante mucho tiempo sin la otra. Pero aunque la expansión de la técnica y la de la economía estén entrelazadas y sean inseparables, ninguna ley simple de la naturaleza establece que la técnica sea la causa del crecimiento económico o que éste impulse el progreso técnico. Un aumento del volumen de los mercados puede resultar en una división más eficaz del trabajo y de la especialización sin introducir novedades técnicas apreciables. La clave está en la interacción entre las personas, las instituciones económicas, el auge de los mercados y el desarrollo de la técnica.

Tras la segunda guerra mundial, muchos intelectuales e ideólogos creían que el futuro de las naciones había que buscarlo en el socialismo o, al menos, en alguna otra modalidad de economía planificada. Las ya veteranas instituciones mercantiles de Occidente eran tachadas además de asociación con el colonialismo. En consecuencia, sólo un puñado de países en vías de desarrollo escogió el modelo occidental, y aun los gobiernos de esos países tomaban una parte más activa en los asuntos económicos que lo que Adam Smith habría prescrito. Ese puñado de países comprendía a Taiwán, Corea del Sur, Hong-Kong y Singapur, que constituyen algunos de los grandes éxitos económicos de este siglo.

Nadie puede garantizar el triunfo a otros países que opten por una vía de actuación política similar. Pero pasando revista a los 40 años de experiencia de regímenes socialistas y frentepopulistas de la Europa del Este, de Asia, de Africa y de Sudamérica, creemos que se requiere una gran dosis de fe contra la experiencia para augurar muchas posibilidades de éxito si no se sigue este camino.

La ciencia, y no ya sólo la occidental, está desplazando la frontera del conocimiento a un ritmo al parecer acelerado. Y como el conocimiento crea recursos económicos y aumenta, por lo general, en proporciones exponenciales, los futuros avances en el bienestar humano puede que sean, por lo menos, tan sorprendentes como los de los últimos 250 años. La ciencia puede desempeñar también un papel mucho más importante en la resolución de problemas de tan hondas raíces como el de la contaminación ambiental y el del aumento de la población, pero sólo abordándolos en el contexto de unas instituciones y unos incentivos personales más eficientes. Las revelaciones que llegan ahora de la Europa del Este indican que allí la contaminación industrial ha sido mucho peor que en Occidente, donde respira la economía de mercado libre.

Dadas las fundadas perspectivas de que los mercados y el comercio internacional continuarán en auge, quizás el desarrollo de Occidente vaya ensanchando la sima que separa hoy a los países que siguen su ejemplo de los que se empecinan en tomar otra vía. Las diferencias económicas están ya originando graves presiones políticas y sociales en los países en vías de desarrollo, sobre todo desde que algunos de ellos han demostrado que las brechas se pueden cerrar. A menos que esas naciones emprendan una acción eficaz, las consecuencias podrían ser aún menos satisfactorias que las que hemos presenciado durante los últimos 40 años.

BIBLIOGRAFIA COMPLEMENTARIA

UNBOUND PROMETHEUS: TECHNOLOGICAL CHANGE AND INDUSTRIAL DEVELOPMENT IN WESTERN EUROPE FROM 1750 TO THE PRESENT. David Landes. Cambridge University Press, 1969.

PHASES OF CAPITALIST DEVELOPMENT. Angus Maddison. Oxford University Press, 1982.

INSIDE THE BLACK BOX: TECHNOLOGY AND ECONOMICS. Nathan Rosenberg. Cambridge University Press, 1983.

HOW THE WEST GREW RICH: THE ECONOMIC TRANSFORMATION OF THE INDUSTRIAL WORLD. Nathan Rosenberg and L. E. Birdzell, Jr., Basic Books, 1986.

THE EUROPEAN MIRACLE: ENVIRONMENTS, ECONOMIES, AND GEOPOLITICS IN THE HISTORY OF EUROPE AND ASIA. Segunda edición; dirigido por E. L. Jones. Cambridge University Press, 1987.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Agujeros negros en los centros galácticos

Algunos agujeros alimentan cuásares brillantes; la mayoría parece dormir en galaxias de calma relativa, como la nuestra. El conocimiento de estos objetos habrá de ayudar a reconstruir la primitiva historia del universo

Martin J. Rees

Las galaxias son los componentes básicos del cosmos. La luz de la mayoría de galaxias representa el resultado conjunto de sus decenas o centenares de miles de millones de estrellas que las forman. Pero desde hace más de 25 años, los astrónomos saben que algunas galaxias también tienen un núcleo central brillante y compacto cuya emisión no se debe a estrellas normales. Los casos más extremos de los núcleos galácticos activos, que así se llaman, son los cuásares, objetos no mayores que el sistema solar cuya radiación total supera la de 100.000 millones de estrellas. Los cuásares parecen representar una etapa particularmente agitada del desarrollo de algunas galaxias. Se van acumulando datos que apuntan que muchas jóvenes galaxias, quizá la mayoría, experimentan una fase en la que se comportan como un cuásar o tienen una actividad parecida.

Los astrónomos suelen coincidir en señalar la gravedad como fuente de energía de los núcleos galácticos activos. Su brillo intensísimo y pequeño tamaño implican que la fuente gravitatoria debe ser, aunque masiva, extremadamente compacta. El candidato con mejor perfil para servir de motor central de los cuásares es un agujero negro, un cuerpo colapsado cuya gravedad es tal que nada, ni siquiera la luz, puede escapar del mismo.

MARTIN J. REES dirige el Instituto de Astronomía de la Universidad de Cambridge, donde recibió su formación académica. Rees ha sido profesor del Instituto de Tecnología de California, la Universidad de Harvard e Instituto de Estudios Avanzados de Princeton. Le interesa la investigación en la estructura y evolución de las galaxias.

El descubrimiento de agujeros negros en los centros galácticos, ya excitante por derecho propio, podría incidir en las concepciones actuales sobre la evolución del universo. Los cuásares aparecieron cuando el universo no había cumplido sus mil millones de años de edad, lo que indica que algunas galaxias habían desarrollado regiones centrales densas. La temprana aparición de los cuásares descarta muchos modelos cosmológicos que predicen que la formación de las galaxias requeriría miles de millones de años y no deja de plantearle problemas al modelo en boga de la materia oscura fría. Las mediciones recientes de la radiación cósmica de fondo acaban de complicar el rompecabezas. Los teóricos piensan, en su mayoría, que las galaxias se formaron a partir de variaciones de densidad en el universo recién nacido. Pero las mediciones de la radiación de fondo indican que las variaciones serían tan minúsculas que se hace difícil explicar cómo pudieron producir las estructuras observadas actualmente. Aparte de su importancia cosmológica, la detección de agujeros negros dotados de gran masa podría también aclarar algunas predicciones de la teoría einsteiniana de la relatividad general.

De acuerdo con esa teoría, un agujero negro galáctico podría denunciar su presencia. El remolino de gas en torno al agujero se acelera hasta casi la velocidad de la luz. El rozamiento calienta el gas, y éste se convierte en un plasma eléctricamente conductor. La gravedad del agujero negro estruja los campos magnéticos que se forman en el plasma, generando calor adicional. Materia así calentada y densa radiaría intensamente. Por este procedimiento, el agujero negro convierte materia en energía con un rendimiento 10 veces superior al de las reacciones nucleares que crean el brillo de las estrellas.

No debiera sorprendernos la presencia de agujeros negros en los centros de las galaxias. Una vez que éstas empezaron a constituirse con la materia primordial, las estrellas o las nubes de gas de las densísimas regiones centrales de las galaxias podrían haberse agregado y colapsado en un agujero negro. La materia adicional que fuera cayendo provocaría que el agujero negro se "encendiera" en forma de cuásar. La energía emitida por el cuásar podría, a su vez, condicionar la formación posterior de la galaxia.

La masa que se acumula en el centro de una galaxia elíptica podría ser proporcional a la masa total de la galaxia. Para galaxias espirales en "disco", la Vía Láctea es una de ellas, el agujero podría guardar relación, no con la masa total de la galaxia, sino con la masa del denso bulbo central de estrellas, que es mucho menor que una galaxia elíptica normal. Los cuásares más luminosos están asociados probablemente con los agujeros negros portadores de mayor masa, los de las galaxias elípticas. Las galaxias espirales que sólo contienen agujeros negros de masa moderada nunca podrían haber sido cuásares potentes.

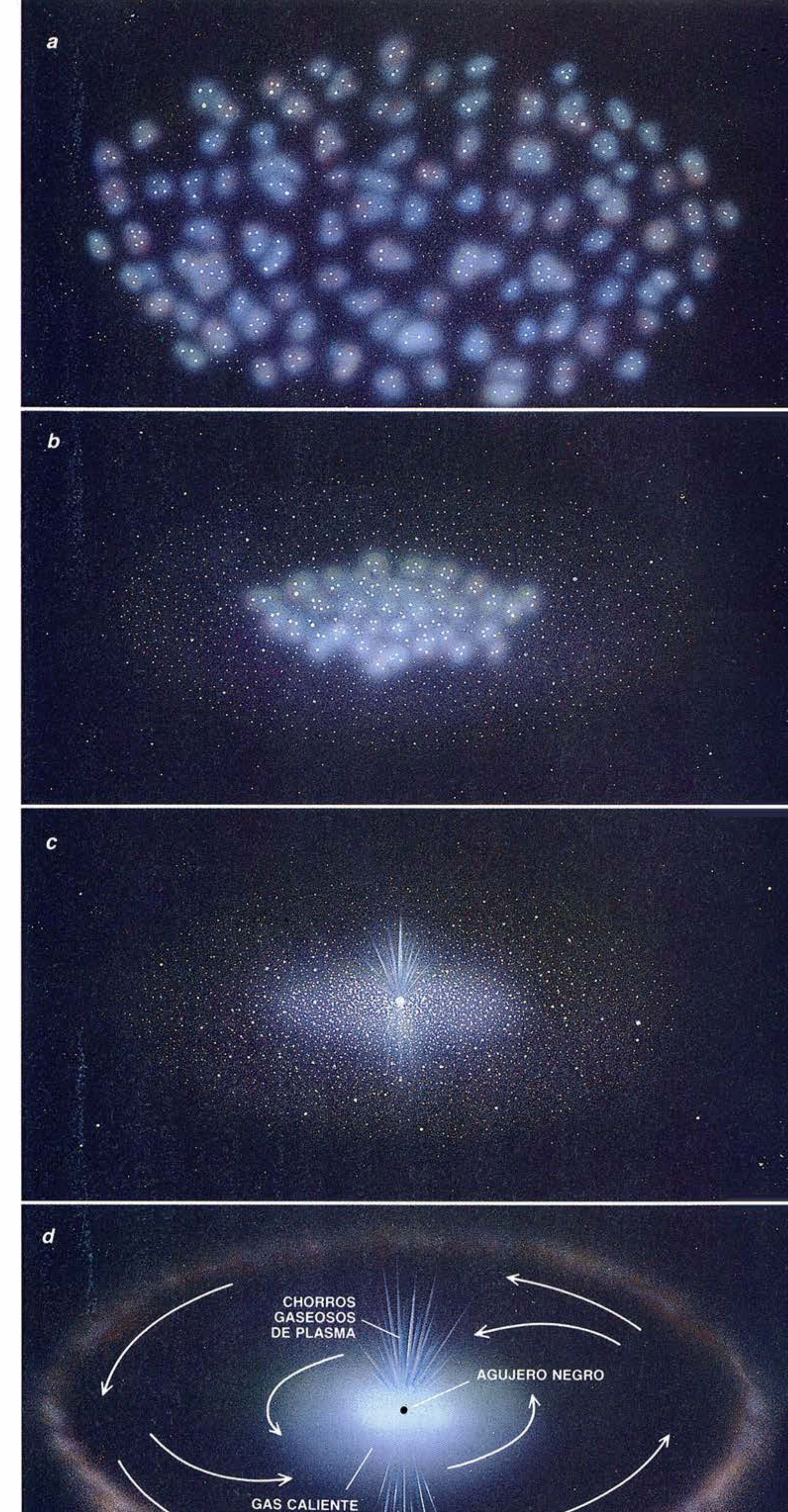
No todas las galaxias elípticas alojan, por necesidad, agujeros negros. Más. Aun cuando un agujero negro habitara un tiempo en una galaxia, podría no estar ya allí. Hay astrofísicos que opinan que muchas galaxias elípticas nacen de la fusión de dos o más galaxias: las órbitas de sus estrellas se mezclan y el amasijo de estrellas resultante adquiere la figura de una elipse. Si las galaxias originales poseían en su centro agujeros negros masivos, los dos juntos describirán espirales hacia el centro de la galaxia unida y se estabilizarán en una órbita

mutua. Los dos agujeros negros se irían aproximando más estrechamente, habida cuenta de que, por la teoría de la relatividad, la energía orbital se radiaría en forma de ondas gravitatorias. Una potente explosión de radiación gravitatoria emitida durante la coalescencia final crearía un retroceso que enviaría violentamente el agujero negro resultante de la fusión hacia el espacio intergaláctico. En principio, las sondas instaladas en el espacio podrían detectar estas potentes ondas de gravedad y confirmar así la validez de las ideas de Einstein.

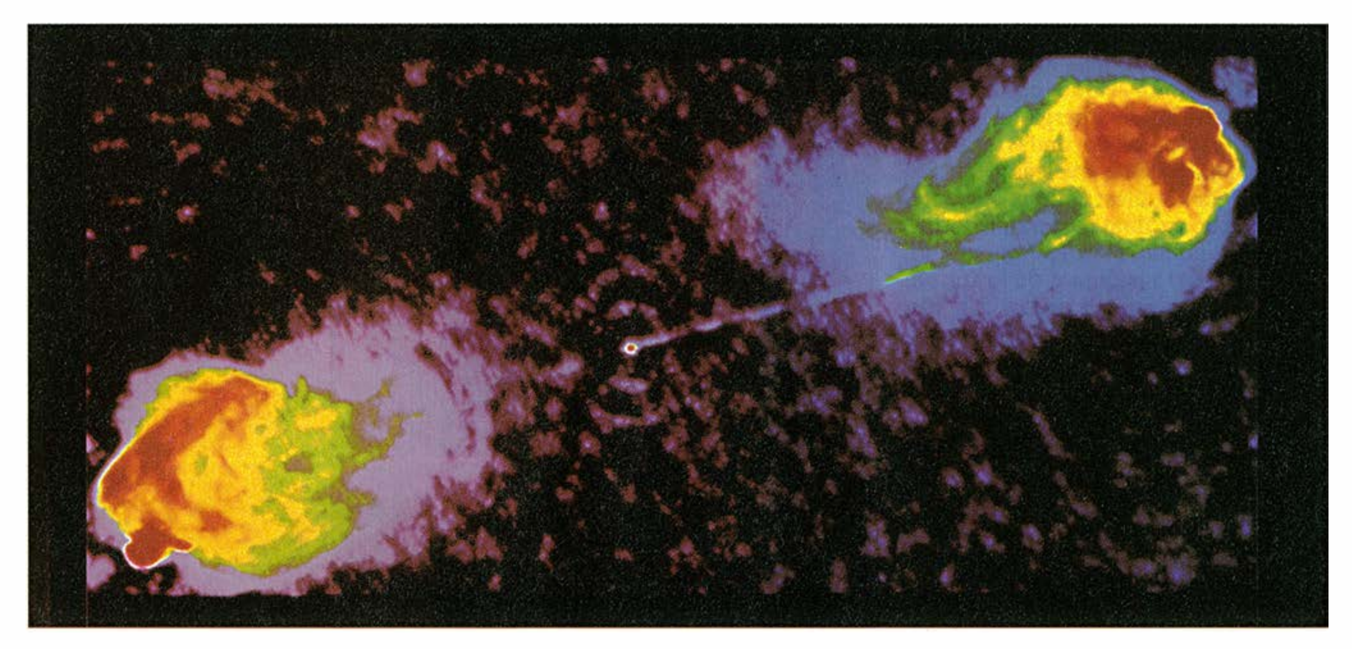
Pero en todo esto nos movemos, por desgracia, en el terreno de la especulación. A los astrofísicos les gustaría conocer los detalles del comportamiento de los cuásares y demás núcleos galácticos activos, su número y cuanto les sucede cuando se desvanecen. Una mejor comprensión de estos objetos energéticos, y enigmáticos, ayudaría a entender el proceso de formación y evolución de las galaxias.

Resulta difícil estudiar los cuásares, alejados como están en el espacio y en el tiempo. Se nos ofrecen tal cual eran hace miles de millones de años, cuando la luz que ahora llega a la Tierra empezó su viaje.

Los astrónomos miden las grandes distancias cósmicas por medio de los corrimientos hacia el rojo, la cuantía en que se ha estirado o "enrojecido" la luz de un objeto. En virtud de la expansión global del universo, los cuerpos se alejan unos de otros (excepto en entornos locales); si consideramos dos cuerpos, éstos se alejan entre sí con una velocidad proporcional a la distancia entre los mismos. La imagen de un globo pintado con topos se nos viene de inmediato a la mente: cuando se infla el globo, cualquier punto de su superficie se aleja de los demás. Cuanto mayor sea la distancia entre dos topos, más deprisa los separa el estiramiento del globo. No se puede señalar un punto como centro de la superficie del globo, de la misma manera que no se puede decir que ninguna galaxia ocupa el centro del universo en expansión.



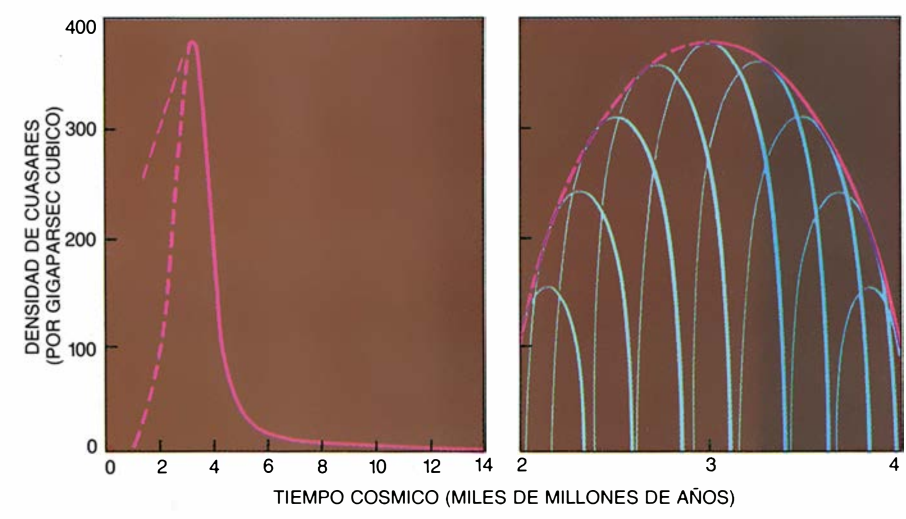
1. NACIMIENTO DE LAS GALAXIAS y su presumible relación con la formación de un agujero negro. Según una teoría evolucionista, una gran nube de gas se colapsa bajo su propia gravedad (*a*) y forma un disco aplanado con una zona central condensada, en la que se congregan muchas estrellas (*b*). La parte de la galaxia más densa e interna sufre un ulterior colapso, resultando un agujero negro; el material que lo rodea reluce cuando se precipita en el agujero (*c*). El tamaño de este último depende del tamaño del bulbo central de la galaxia. Los agujeros mayores alimentarían los cuásares brillantes (*d*).



2. GALAXIA DEL CISNE A, la radiofuente más intensa del cielo; radia principalmente desde dos lóbulos de plasma (gas ionizado) que tienen un tamaño de cientos de miles de años-luz. Los lóbulos están alimentados, se supone, por chorros calientes que salen arrojados cuando cae materia gaseosa sobre un gran agujero negro del centro de la galaxia. La energía de los lóbulos equivale a millones de masas solares; del tamaño y la estructura de los lóbulos se infiere que Cisne A se mantuvo en actividad durante decenas de millones de años. (Observatorio Nacional de Radioastronomía.)

Los cuásares más remotos presentan grandes corrimientos hacia el rojo, hasta el extremo de que las longitudes de onda de su radiación se han estirado más de cinco veces entre la emisión y la recepción. La luz visible de estos cuásares que alcanza la Tierra se emitió en su origen en forma de radiación ultravioleta, de longitud de onda mucho más corta.

Podemos asociar el factor de estiramiento, o de desplazamiento hacia el rojo, con la razón entre el tamaño actual del universo y el que tenía cuando se emitió la luz. El desplazamiento hacia el rojo de un cuásar mide cuán grande, y por tanto cuán viejo, era el universo cuando el cuásar emitió la luz. Al observar los cuásares, los astrónomos están viajando por la historia cósmica remontándose hasta la era en que el universo tenía un tamaño menor que una quinta parte del actual. El modelo cosmológico más en boga defiende que el universo tiene unos 13.000 millones de años y que estos cuásares remotos brillaban cuando el universo tenía sólo mil millones de años.



3. POBLACION DE CUASARES en el universo primitivo. Fue mucho mayor y estuvo más concentrada en el intervalo de dos a cuatro mil millones de años después de la gran explosión que en la actualidad (*izquierda*). En épocas anteriores, la población mengua, aunque los datos son fragmentarios. Las indicaciones astrofísicas sugieren que la concentración lentamente variable de cuásares oculta en realidad muchas generaciones superpuestas de objetos de vida más o menos corta (*derecha*).

El análisis de varios miles de cuásares revela que la población ha cambiado radicalmente con el tiempo. Cuando el universo tenía unos dos mil millones de años, los cuásares abundaban mucho más que ahora. La actividad de los cuásares conoció su apogeo en la época en que la escala del universo correspondería a un tercio de la actual, unos dos o tres miles de millones de años después de la gran explosión. Por lo que parece, los cuásares más viejos son escasos, no sólo porque se hallan más alejados y son, por tanto, más débiles, sino también debido a que la población de cuásares decrece para corrimientos hacia el rojo correspondientes a épocas anteriores a los dos mil millones de años tras la gran explosión.

La mayoría de los cuásares que otrora se mostraron activos, no existen ya. El cuásar brillante más cercano, llamado 3C273, se encuentra a dos mil millones de años-luz de la Tierra y tiene un brillo aparente de magnitud 14, mil veces por debajo del umbral de la visión humana. Durante la "era cuásar", hace unos 11.000 millones de años, el cuásar más cercano se hubiera hallado a sólo unos 25 millones de años-luz y hubiera alcanzado la luminosidad de una estrella de magnitud cuatro, fácilmente visible por el ojo humano. Que los cuásares estuvieran entonces más cerca resulta obvio, por la sencilla razón de que todo el universo era más pequeño y más denso. Pero también la relación de cuásar a galaxia multiplicaba en-

tonces por 1000 la razón de frecuencia actual.

Es sorprendente el poco tiempo que media desde la gran explosión hasta la aparición de los primeros cuásares. Los astrónomos han encontrado algunos cuásares cuya luz se ha estirado en un factor de casi seis, señal que revela su presencia cuando el universo no tenía ni mil millones de años. Estos cuásares viejos y alejados imponen serias limitaciones a las teorías sobre la formación de galaxias.

Presumiblemente, la actividad de los cuásares no pudo empezar por lo menos hasta que algunas galaxias se hubieran condensado y desarrollado regiones centrales concentradas. Sólo entonces el potencial gravitacional del centro habría adquirido valor suficiente para provocar un colapso gravitacional desenfrenado, la formación de un agujero negro y el nacimiento de un cuásar. El descubrimiento de que los cuásares aparecieron en fase tan precoz de la historia del cosmos pone en aprietos algunas teorías de formación de galaxias, a tenor de las cuales las galaxias no empezaron ni siquiera a coalescer hasta mucho más tarde.

Nos desconcierta la drástica subida y caída de la población de cuásares y, en realidad, de todas las clases de núcleos galácticos activos. La actividad cuásar es nota característica de galaxias muy jóvenes, debido quizás a que, conforme evoluciona la galaxia, va quedando atrapada en las estrellas una porción creciente de su gas, dejando menos combustible disponible para los cuásares.

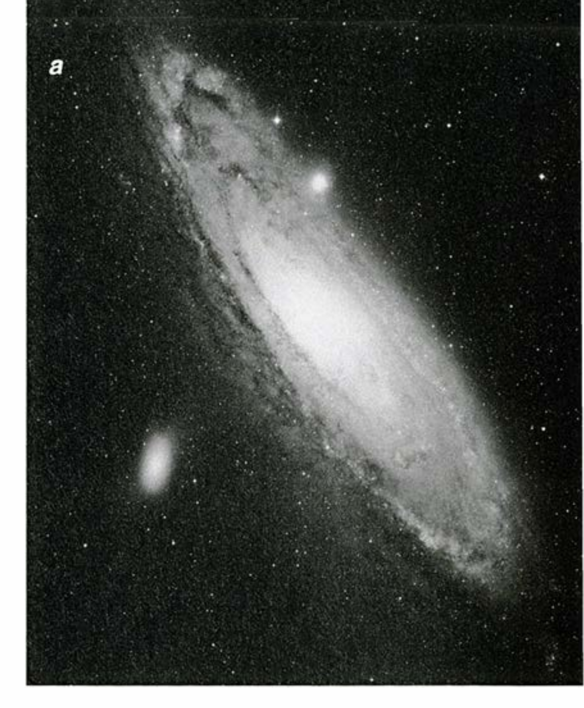
La observación, por sí sola, no puede revelarnos si los cuásares constituyen una etapa común en la vida de todas las galaxias jóvenes o son una aberración, eso sí muy aparatosa, de algunas insólitas. Para contestar esta pregunta hace falta saber cuánto tiempo vive un cuásar normal.

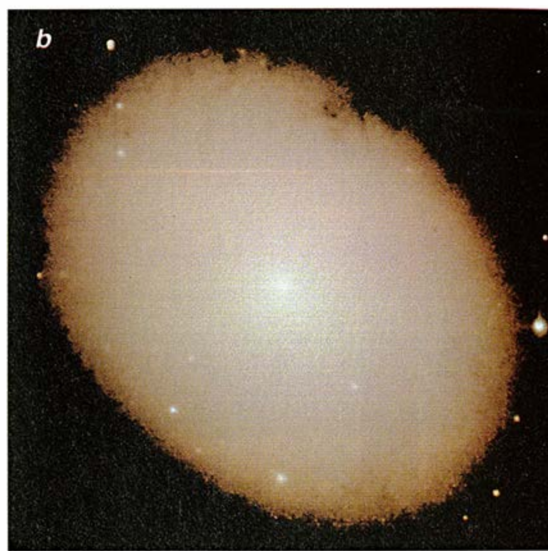
La proporción actual de cuásares no llega a uno por cada 100.000 galaxias. Durante la era de los cuásares, hace 11.000 millones de años, la razón era de uno a cien, más o menos. Es decir, sólo una galaxia de cada 100 ha pasado alguna vez por un período de actividad cuásar; sólo, pues, el 1 por ciento de todas las galaxias actuales, aproximadamente, deben albergar remanentes. Pero hay una posibilidad alternativa: muchas generaciones de cuásares de vida corta podrían nacer, desarrollarse y morir durante el período en que la población total aumenta y declina. (Muchas generaciones de habitantes pueden vivir y morir mientras la población de una ciudad crece y mengua en una escala de siglos.) Puede que cada cuásar individual sólo brille por poco tiempo, posiblemente mucho menos de mil millones de años. Si es así, los remanentes de los cuásares abundarían mucho más que si sólo hubiera existido una generación.

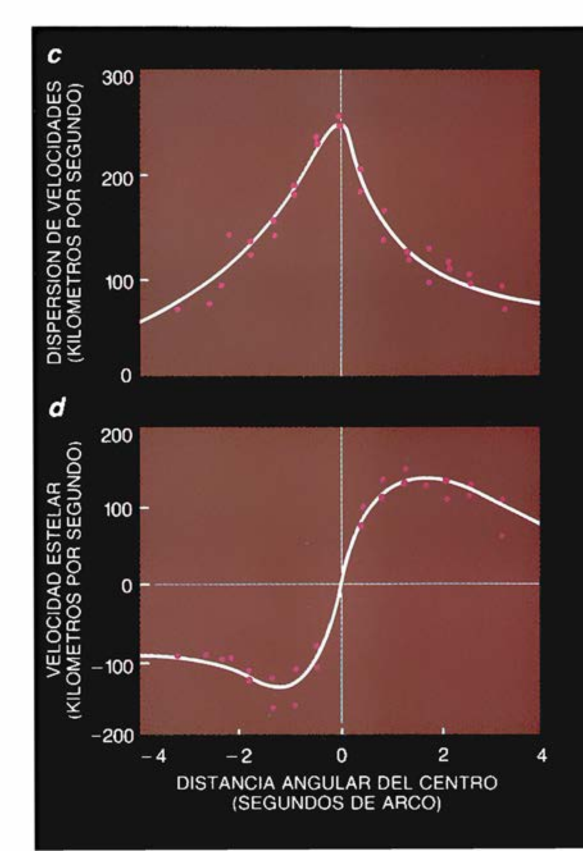
Para calcular el tiempo que permanece activo un cuásar podemos estimar su masa. Los más potentes emiten tanta energía que deben engullir varias decenas de masas solares cada año. Si estos cuásares permanecieran activos mil millones de años acabarían pesando decenas de miles de millones de masas solares. Hasta los menos luminosos, los más corrientes, pesarían más de mil millones de masas solares. Pero si se hubieran sucedido muchas generaciones de cuásares de vida corta, ninguno de ellos habría tenido tiempo de acumular tanta masa.

Existe una buena razón física para pensar que los cuásares deben pesar, por lo menos, 100 millones de soles, con independencia del tiempo que vivan. La radiación ejerce una ligera presión. La presión de una radiación intensa expelería toda la materia de la vecindad de un cuásar, a menos que la gravedad la contrarrestara con una fuerza de intensidad suficiente. Si los cuásares se alimentan por acreción, atrayendo el gas de su entorno, la cuantía de su masa debe ser tal que permita a la gravedad vencer la presión de radiación. Por tanto, la luminosidad de un cuásar establece un límite inferior para la masa. Los cuásares normales pesarán por lo menos 100 millones de soles; los más potentes, mil millones de soles como mínimo.

Para estimar la masa de un grupúsculo de cuásares que pertenecen a la categoría de radiofuentes podemos echar mano de otro método. La emisión de radio surge, en general, de dos grandes lóbulos de plasma excitados por el cuásar. La energía almacenada en los lóbulos equivale a la masa de millones de soles; suponiendo niveles de rendimiento plausibles, los cuásares deben haber acumulado una masa de 100 millones de soles, como mínimo, para liberar tanta energía. La estructura de los lóbulos radioemisores asociados a estos cuásares también permite deducir que la







4. GALAXIA DE ANDROMEDA (*a*). Posee masa compacta en su centro, un agujero negro quizá. La ampliación de la región del núcleo (*b*) muestra la gran concentración de estrellas que reina allí. La dispersión de velocidades estelares (*c*) y las velocidades orbitales (*d*) aumentan de forma manifiesta en el centro, señal de la intensa atracción de un objeto masivo invisible.

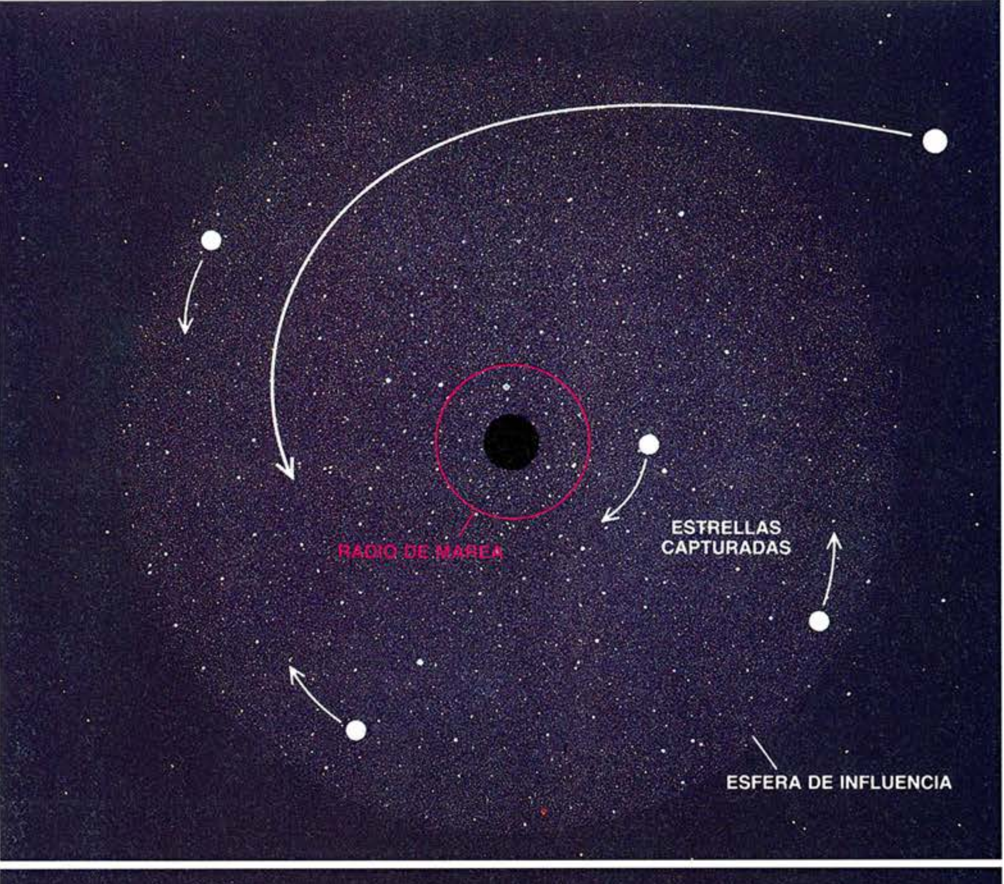
actividad de éstos se ha desarrollado, a lo largo de toda su vida, sin solución de continuidad.

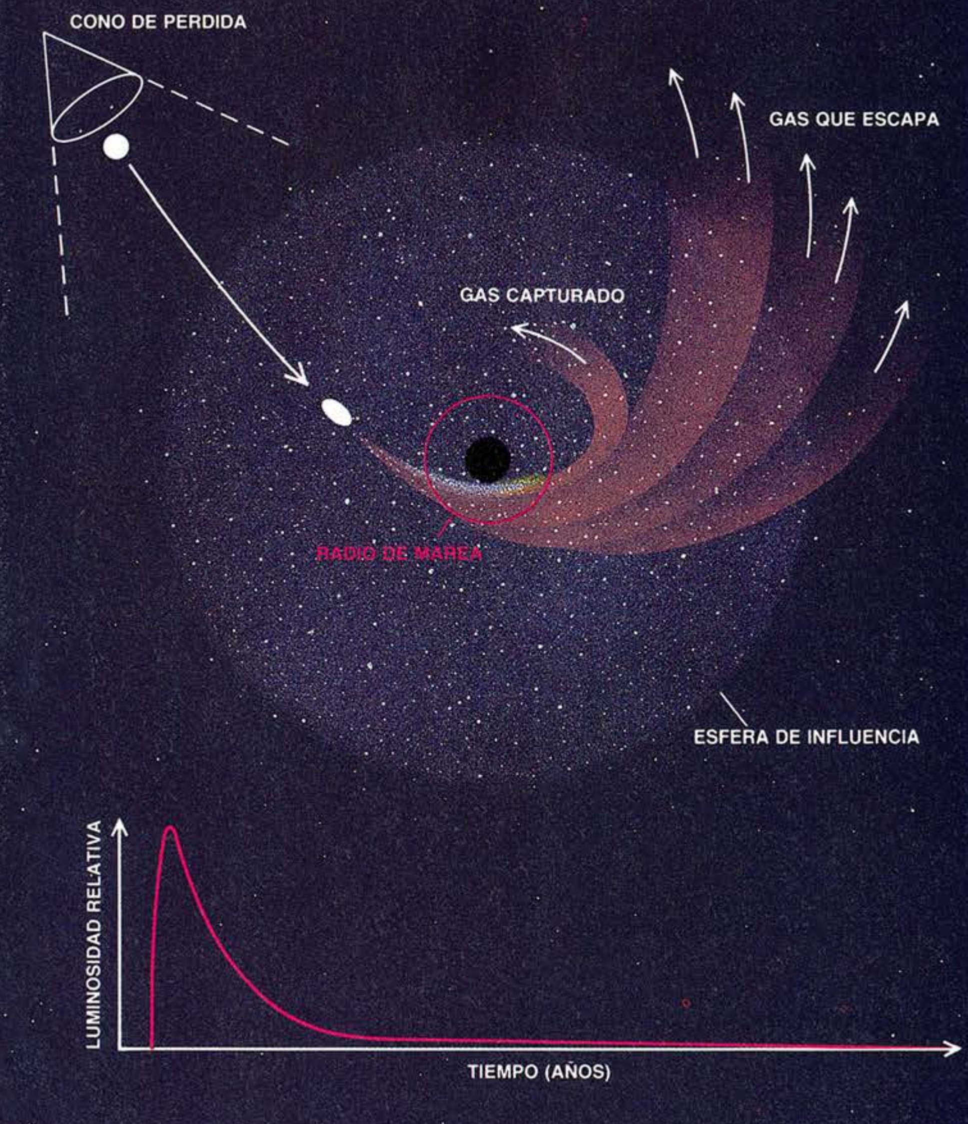
Por desgracia, no se sabe con cuánta masa nacen los cuásares ni cuánta acumulan antes de que se agote su combustible. Los estudios espectroscópicos revelan que la presión de radiación afecta decisivamente a los cuásares; implica ello que no recaban la masa que se requeriría para que la gravedad desbordara la fuerza de la radiación. Con este razonamiento resulta coherente admitir que un cuásar normal no brilla durante más de 50 millones de años. La relativa brevedad de sus vidas significa que muchas galaxias, quizá todas, experimentan en su juventud una fase en la que se comportan como un cuásar o muestran una actividad parecida.

Estas conjeturas, ninguna de ellas muy precisa, son las únicas con que cuentan los astrónomos acerca de las masas y tiempos de vida de los cuásares activos. Las estimaciones teóricas de la masa pueden refinarse calculándola por separado para subclases que tengan, por ejemplo, distintas luminosidades y características de radiación; pero persisten grandes incertidumbres acerca de las vidas medias deducidas.

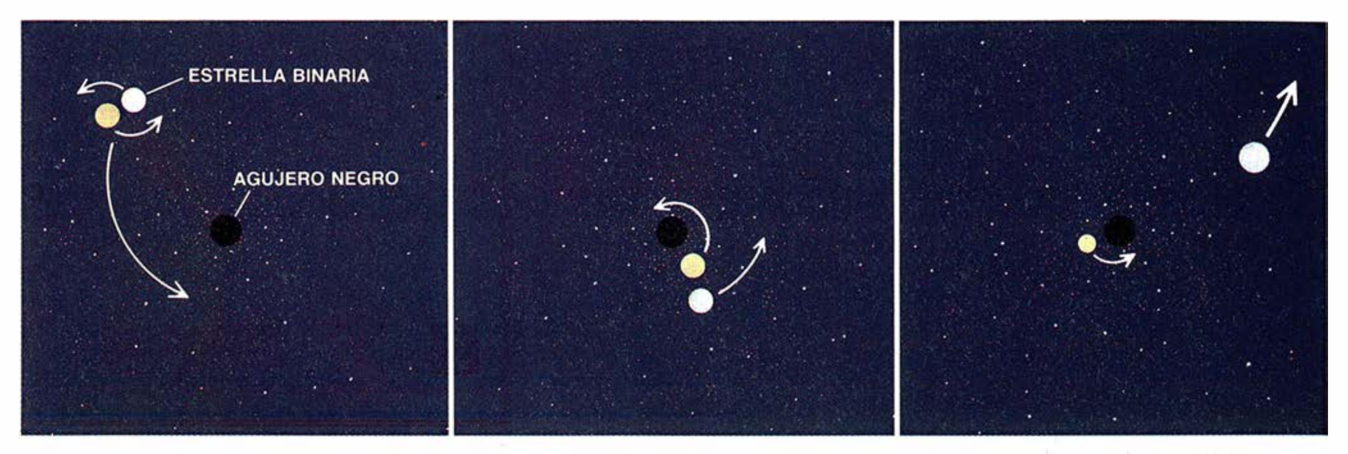
Otra vía de ampliar el conocimiento que tenemos de la naturaleza de los cuásares y de otros núcleos galácticos activos es la que nos lleva a buscar sus cadáveres: podría haber agujeros negros inactivos en los centros de muchas galaxias cercanas. Y en esa línea, medir directamente las masas y concentración de tales objetos.

Un agujero negro masivo sólo se comportará como cuásar mientras se alimente capturando gas y polvo de su entorno. En un vacío, no emitirá ninguna radiación observable. Pero aun entonces seguirá ejerciendo una influencia gravitacional detectable sobre los cuerpos que lo rodeen. Un agujero negro grande en el núcleo de una galaxia atraería las estrellas hacia sí y provocaría la acumulación de las mismas en órbitas cercanas al centro. Las estrellas próximas al agujero negro describirían órbitas celerísimas.





5. EL AGUJERO NEGRO CAPTURA todas las estrellas incluidas en un radio crítico con su atracción gravitacional, produciendo un cúmulo característico de estrellas que se mueven muy deprisa (*arriba*). Las estrellas que se aproximan al agujero con ángulos pequeños penetran en el radio de marea, en cuyo interior la gravedad del agujero desgarra la estrella (*en medio*). Parte del gas de la estrella cae en el agujero, creando entonces una breve fulguración. La mayor parte del tiempo, el agujero se encontraría inactivo.



6. ACONTECE EL EFECTO DE LA HONDA GRAVITATORIA cuando una estrella binaria se aproxima mucho a un agujero negro. Una estrella puede ser capturada mientras que la otra sale despedida a unos 10.000 km por segundo. Los encuentros normales entre estrellas no acaban así.

Toda estrella de una galaxia responde al campo gravitacional acumulado del resto de la materia de la galaxia. Las estrellas orbitan con velocidades que equilibran el efecto disgregador de su energía cinética con la atracción de la gravedad que intenta llevarlas hacia el centro. Las estrellas pueden quedar capturadas en órbitas pequeñas y rápidas si se aproximan a la vecindad de un agujero negro central.

El radio de la "esfera de influencia", que sobre las estrellas del entorno ejerce un agujero negro, es directamente proporcional a su masa y depende inversamente del cuadrado de la velocidad de la estrella. En otras palabras, cuanto más deprisa se mueva una estrella, menos le afectará la gravedad del agujero negro. La esfera de influencia es millones de veces mayor que el propio agujero negro; a pesar de ello, poseerá sólo una anchura de decenas de años-luz. En las mismas galaxias próximas, una región tan pequeña abarcaría un ángulo de escasos segundos de arco (un segundo de arco es 1/3600 de grado).

Un signo que delataría la presencia de un agujero negro inactivo en el corazón de una galaxia sería un pico agudo en la curva de luz de la galaxia, resultante de la concentración de estrellas en la esfera de influencia del agujero negro. Algunos investigadores han buscado tales picos, pero en los resultados siempre queda la incertidumbre de que otras fuentes distintas de las estrellas, por ejemplo las nubes de gas luminosas, contribuyan a la luz del centro.

Un indicio menos ambiguo de un agujero negro sería la presencia de estrellas que se movieran con velocidad anómala, por excesiva, muy cerca del centro de la galaxia. Ahora bien, no pueden detectarse estrellas individuales de brillo ordinario en otras galaxias, y menos en los centros, que están abarrotados. Los detalles más finos que pueden resolver los telescopios ópticos desde el suelo tienen tamaños angulares de medio segundo de arco aproximadamente. Este tamaño angular corresponde a una región de cinco años-luz de ancho en la galaxia de Andrómeda, que está a una distancia de unos dos millones de años-luz de la Vía Láctea y es la más cercana de las grandes galaxias espirales parecidas a la nuestra; una región de este tamaño alberga decenas de miles de estrellas.

No necesitamos, por suerte, observar estrellas por separado para extraer las características de interés de su movimiento. Los espectros ópticos de la luz combinada de las estrellas de la región central pueden proporcionar la información deseada. Siempre aparecen unas marcas oscuras características (las líneas de absorción) a determinadas longitudes de onda de la emisión luminosa estelar. En un cúmulo orbitante, unas estrellas avanzan hacia la Tierra; otras se alejan. La luz de las que se alejan se estira (se corre hacia el rojo); la luz de las que se acercan se comprime (se corre hacia el azul). Las líneas del espectro total del cúmulo quedan difuminadas por los desplazamientos superpuestos hacia el rojo y hacia el azul. El ensanchamiento de las líneas indica el rango de velocidades de las estrellas. Las líneas muy anchas señalarían la presencia de un objeto muy denso, quizás un agujero negro.

En 1979, Peter J. Young y Wallace L. W. Sargent, del Instituto de Tecnología de California, y sus colaboradores, encontraron datos espectrales de una masa compacta de tres mil millones de soles en el núcleo de la galaxia gigante elíptica M87. Esta galaxia es una radiofuente intensa y presenta un chorro insólito de materia luminosa que parece partir de su núcleo; ello indica una actividad central, si bien con un nivel de energía muy por debajo del que posee el cuásar. Si buscamos agujeros negros centrales, M87 es una buena candidata, aunque una masa de tres mil millones de soles, que implica que M87 albergó un cuásar brillante, les parece sorprendentemente grande a algunos teóricos. A la distancia de M87, 35 millones de años-luz de la Tierra, los efectos de un agujero negro de menos masa no podrían detectarse con los instrumentos habituales.

Durante la última década la credibilidad de las observaciones de M87 fue objeto de debate. Discutíase, sobre todo, si debía haber masa oscura en el centro para justificar las altas velocidades estelares medidas, o si la concentración de estrellas aportaba masa suficiente para explicar los movimientos. Un problema, y no pequeño, lo ha constituido la dificultad de distinguir entre la luz de las estrellas cercanas al núcleo y la luz no estelar emitida por lo que alimenta el chorro luminoso, sea lo que fuere.

El análisis de los vecinos galácticos inmediatos a la Vía Láctea ha ofrecido resultados más nítidos. Por hallarse estas galaxias a una distancia mucho menor que M87 (10 veces más cerca), podría advertirse una concentración central de masa de 10 o 100 millones de soles. A diferencia de M87, las galaxias cercanas no muestran signos de actividad, ni siquiera a niveles bajos. Según se ha indicado antes, el agujero negro no revelaría su presencia si no recibe materia.

Los astrónomos ópticos han encontrado pruebas indirectas de masas centrales compactas en varias galaxias cercanas. El caso más interesante y convincente tiene que ver con Andrómeda. En los años sesenta, Martin Schwarzschild y sus colaboradores, de la Universidad de Princeton, construyeron el *Stratoscope*, pequeño teles-

copio que fue lanzado en globo a mucha altura a fin de evitar el efecto difuminador de la baja atmósfera. En su último vuelo, en 1971, descubrieron que las estrellas que se encuentran a pocos años-luz del centro de la galaxia de Andrómeda formaban un disco aplanado. En los últimos años Alan M. Dressler, del Instituto Carnegie, Douglas O. Richstone, de la Universidad de Michigan, y John Kormendy, de la Universidad de Hawai, han sacado los espectros para analizar los movimientos de esas estrellas. Han encontrado que las velocidades aumentan súbitamente al acercarse al centro. Además, el sistema estelar aplanado gira de una manera que indica que la masa de su centro está extremadamente concentrada. Se cifra entre 30 y 70 millones de soles la masa del objeto central.

Otras galaxias cercanas muestran comportamientos similares. John L. Tonry, actualmente en el Instituto de Tecnología de Massachusetts, demostró, en 1984, que la pequeña galaxia M32 (vecina de Andrómeda) parecía albergar un objeto central oscuro de unos cinco millones de masas solares. La galaxia Sombrero (nombre que atinadamente describe su forma) fue estudiada por B. J. Jarvis y P. Dubath, en el Observatorio Meridional Europeo, y por Kormendy con el telescopio Canadá-Francia-Hawai de Mauna Kea. La galaxia Sombrero parece tener un núcleo estelar en rápida rotación que rodea una masa central concentrada de casi mil millones de masas solares.

Ahora bien, ¿son necesariamente agujeros negros estas masas oscuras y compactas? La falta de un gran pico central en la distribución de luz de Andrómeda implica que el objeto central, sea lo que fuere, emite muy poca luz en relación a su masa. ¿Podría tratarse de una población insólita de estrellas débiles, concentrada cerca del centro? Un cúmulo abarrotado de pequeños agujeros negros (de entre 10 y 100 masas solares cada uno), de estrellas de neutrones o de enanas marrones (objetos cuya masa no llega a la décima parte de la solar, lo cual impide que puedan brillar) podría contener una masa considerable a pesar de dar poca luz. El motivo de que tales objetos se acumulasen en los diez años-luz más internos de una galaxia, lo mismo que la forma de hacerlo, es algo desconocido, pero no puede descartarse su realidad.

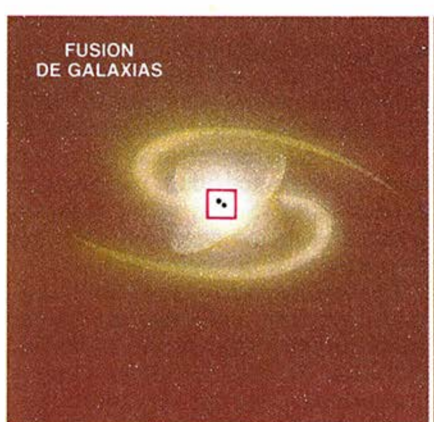
Observaciones más precisas podrían dirimir la cuestión. Si las estrellas inmediatas al centro de la galaxia se mueven más deprisa que las que se hallan a un segundo de arco de su centro (lo que implicaría que la masa oscura está confinada dentro del año-luz más interno, no simplemente en una zona de 10 años-luz) entonces quedarían descartadas muchas de las hipótesis que apelan a las estrellas oscuras. Un cúmulo denso de estrellas de sólo un año-luz de anchura evolucionaría rápidamente debido a los frecuentes encuentros estelares. En poco tiempo las estrellas más pesadas se acumularían en el medio y colapsarían en un cuerpo central, presumiblemente un agujero negro. Carece de verosimilitud que un cúmulo estelar central denso hubiera sobrevivido intacto en una galaxia de 10.000 millones de años.

Si los agujeros negros centrales y grandes son objetos habituales, deberán ser también sumamente tranquilos. Los astrónomos recurren a los agujeros negros para explicar algunas de las fuentes de radiación más luminosas y eficientes del universo. Pero ningún signo hay de elevados niveles de actividad en Andrómeda, la galaxia Sombrero o M32. Las observaciones revelan que Andrómeda no emite más de una diezmilésima de la radiación de un cuásar. En la mayoría de galaxias, la región más interna podría muy bien hallarse exenta de gas.

Incluso un agujero negro inactivo podría ocasionalmente volver a brillar de nuevo si consumiera una estrella de las regiones internas de la galaxia. Cada estrella de la galaxia describe una órbita compleja bajo la influencia gravitacional combinada de todas las demás estrellas y del propio agujero. Poco a poco, las órbitas van cambiando o difuminándose debido al efecto acumulativo de los encuentros con otras estrellas. Los encuentros pueden poner una estrella en una órbita que pase cerca del agujero. De vez en cuando puede caer alguna estrella directamente en el agujero.

Si una estrella se acerca demasiado a un agujero negro puede sufrir las consecuencias. Los cuerpos gaseosos, las estrellas entre ellos, son sensibles a los efectos de marea, el gradiente de la tensión gravitacional en el cuerpo. Si una estrella pasa demasiado cerca de un agujero negro masivo, la diferente atracción gravitatoria que experimentan las partes de la estrella más cercanas y más alejadas al agujero la desgarrarían. Cualquier estrella que entrara dentro del "radio de marea" quedaría destrozada.

En la linde de un agujero negro las fuerzas de marea son más suaves para los agujeros de mayor masa. El radio de una agujero negro aumenta en proporción directa a su masa, mientras que el radio de marea crece sólo con la raíz cúbica de la masa. En el caso de agujeros negros dotados de gran masa, así el de 3000 millones de masas solares cuya existencia en M87 se postula, el radio de marea se halla dentro del horizonte de sucesos, la "superficie" oscura que señala el punto de no retorno del agujero negro.







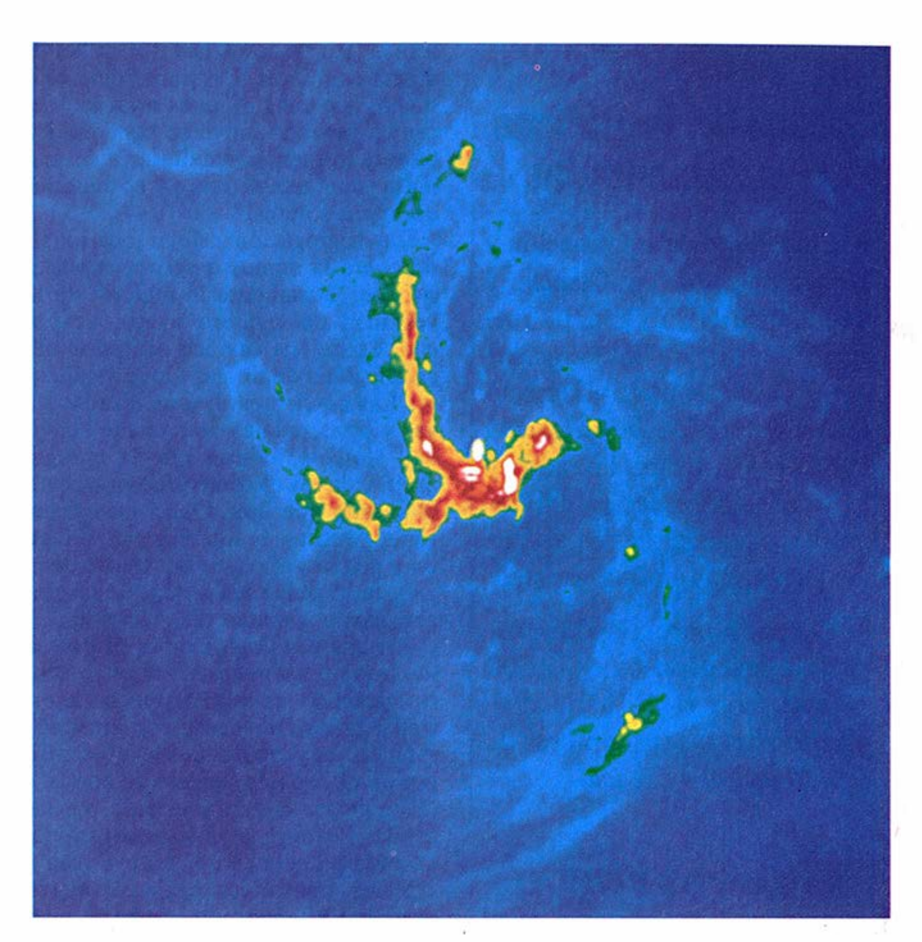
7. CON LA UNION DE DOS GALAXIAS se ponen en órbita mutua los agujeros negros de los dos centros galácticos. Los agujeros emiten radiación gravitacional, que provoca el decaimiento de sus órbitas. Los dos acaban por combinarse en un único objeto. Si sus masas son desiguales, o si los agujeros negros se hallan en rotación, la ráfaga final de radiación gravitacional podría expulsar de la galaxia al agujero que se han unido con otro.

Las estrellas que se aproximen al agujero sólo se destrozarían tras caer irreversiblemente en el interior de su horizonte. La situación cambia cuando se trata de agujeros negros más pequeños, de 5 a 100 millones de masas solares (los observados en las galaxias cercanas), cuyo radio de marea multiplica de 10 a 100 veces el propio agujero. Las estrellas se desgarrarían cuando distaran lo suficiente de esos objetos como para que se escapara la radiación dimanada del encuentro.

Tales sucesos deberían ocurrir aproximadamente una vez en miles de años, no muchos según cálculos de dinámica estelar. El ritmo exacto depende de la estadística de las órbitas estelares y, en particular, de la celeridad con que las nuevas estrellas ocupan las órbitas vacías que pasan cerca del agujero. Cuando se destroza una estrella, la repentina liberación de gas produce inevitablemente alguna radiación. El destello terminal de una estrella destrozada podría ser la señal más clara de la presencia de un agujero negro. La teoría de la relatividad predice la morfología del destello; la observación de un ejemplar constituiría otra prueba más de esta teoría.

Cuando las fuerzas de marea del agujero negro empiezan a competir con la gravedad propia de una estrella, el material de ésta responde de una manera complicada, estirándose según la dirección orbital y contrayéndose según la perpendicular a la órbita, en medio de fuertes agitaciones y convulsiones. La reproducción exacta de este fenómeno es un reto que se sigue resistiendo incluso a las simulaciones por ordenador. Los próximos años, con mejores simulaciones, habrán de revelarnos qué sucede cuando las mareas no destrozan estrellas de distintas clases y cuáles serán los fenómenos característicos de esos episodios que podremos observar. Los cálculos todavía provisionales de mis colegas y míos deben, sin embargo, darnos una idea —sin más pretensiones— de lo esencial del proceso.

Los restos de la estrella se desparraman en un conjunto de órbitas en torno al agujero. La mitad aproximada de la materia permanece ligada, por gravedad, al agujero. El resto se esparce en órbitas hiperbólicas a velocidades de hasta 10.000 kilómetros por segundo. ¿Qué vería un observador si tal suceso tuviera lugar en el centro de la galaxia de Andrómeda? Los efectos más notorios provendrían del material que se arremolina hacia el agujero. El núcleo de la galaxia parecería encenderse con una potencia mucho mayor que la de una supernova, casi comparable al brillo de un cuásar, durante un año más o menos.



8. RADIOCARTOGRAFIA del centro de la Vía Láctea. Muestra brazos de gas perturbados, quizás el resultado de la disgregación de una estrella por un agujero negro. Los movimientos del gas y de las estrellas en la vecindad del centro señalan la presencia de masa oscura, pesada y compacta.

Es difícil calcular qué proporción de esta radiación emerge a través de luz visible y cuánta a través de radiación ultravioleta de alta energía o de rayos X. Los modelos actuales tampoco permiten predecir la tasa de apagamiento del brillo de la fulguración, que depende del tiempo que tarde la estrella en ser digerida. Se trata de un punto importante; a los astrónomos les gustaría saber si la radiación proveniente de la región que rodea el agujero negro se reduce hasta niveles por debajo del límite de detección antes de que se produzca la siguiente disgregación estelar, 1000 o 10.000 años después. Cuanto más tiempo permanezca visible una fulguración, más probable es que alguna se pueda detectar ahora en una galaxia cercana.

Si la mayoría de galaxias albergan agujeros negros, entonces una muestra de las 10.000 galaxias más próximas (hasta el cúmulo de Virgo, a unos 60 millones de años-luz de distancia) debería contener unas cuantas galaxias que estuvieran hoy en el apogeo de una fulguración. Otras galaxias, presumiblemente más numerosas, tendrían que hallarse en un estado tal que los efectos que aún quedaran de la disgregación por marea más reciente todavía fueran discernibles. La vigilancia astronómica organizada para detectar supernovas en otras galaxias se está encauzando hacia la búsqueda de esas fulguraciones. Sería del mayor interés controlar anualmente las 1000 galaxias del cúmulo de Virgo con el fin de descubrir pruebas de un despedazamiento estelar causado por un agujero negro.

Si en la mayoría de galaxias cercanas residen agujeros negros, la Vía Láctea, que es la nuestra, parecería castigada si no abrigara también el suyo. Las complejas estructuras de la vecindad de nuestro centro galáctico incluyen varias fuentes de infrarrojos asimétricamente dispuestas con respecto a una radiofuente intensa y compacta [véase "¿Qué está ocurriendo en el centro de nuestra galaxia?", por Charles H. Townes y Reinhard Genzel; INVESTIGACIÓN Y CIENCIA, junio de 1990]. La fuente está en el centro gravitatorio de la galaxia, o muy cerca del mismo, y tiene un aspecto muy distinto del resto de las radiofuentes de la galaxia. Su emisión se explica mediante la acreción de pequeñas cantidades de gas en un agu-

jero negro de unos cuantos millones de masas solares. Las nubes de gas situadas a menos de unos años-luz del centro parecen moverse con rapidez bajo la influencia gravitatoria de una masa compacta de dos o tres millones de masas solares.

La disgregación estelar ocasional, tal como la hemos descrito, expulsaría energéticamente la materia de la vecindad del agujero negro; de hecho, existe una cavidad central ahuecada y llena de un gas tenue y caliente. Las corrientes desorganizadas de gas y otras señales de reciente actividad violenta indican que la energía producida en el centro fue mayor en el pasado que hoy. Quizá las estructuras gaseosas, braquiales y situadas a unos cuantos años-luz del centro de la galaxia, perfilan los rastros de disgregaciones por marea más recientes.

Un buen indicador de la presencia de un agujero negro en nuestro centro galáctico sería la existencia de estrellas que se movieran con inusitada rapidez, aceleradas desde el centro por el así llamado "efecto de honda". (Este efecto fue aprovechado por las sondas *Voyager*: el tirón gravitacional de cada planeta estudiado ayudaba a acelerar las sondas hacia su nuevo destino.) El efecto de honda no puede impartir una velocidad mayor que la de escape del cuerpo principal, que es la velocidad crítica necesaria para liberarse de la ligadura gravitacional del cuerpo. Las estrellas eyectadas por sistemas estelares dobles o múltiples no pueden, por tanto, alcanzar velocidades que superen los 1000 kilómetros por segundo, la velocidad de escape de una estrella típica.

Cuando una estrella binaria pasa cerca de un agujero negro central dotado de gran masa, una de las estrellas puede ser capturada en una órbita en torno al agujero, mientras que la otra es despedida a altísima velocidad. La estrella que escapa saldría a más de 1000 kilómetros por segundo, según cálculos de Jack G. Hills, del Laboratorio Nacional Los Alamos.

El descubrimiento de tan sólo una estrella que se alejara del centro galáctico con una velocidad mucho mayor que la velocidad de escape estelar normal constituiría una prueba contundente de la existencia de un agujero negro central dotado de gran masa. La densidad de estrellas en el corazón de la Vía Láctea no se conoce bien; pero si son binarias el 1 por ciento de las estrellas, cada 10.000 años una estrella sería catapultada por el agujero negro con una velocidad de unos 4000 kilómetros por segundo. A este ritmo, cada una tardaría unos dos millones de años en recorrer los 30.000 años-luz que separan el Sol del centro. Unas 200 (dos millones dividido por 10.000) de estas estrellas supercéleres deberían estar por tanto alejándose del centro, aunque dentro del radio de la órbita solar en la galaxia. Supuesto que su brillo sea suficiente para detectarlas, tales estrellas deberían, en principio, ser fáciles de reconocer, ya que se moverían rápidamente por el cielo.

Las estrellas supercéleres eyectadas por el agujero negro adquieren su energía cinética a expensas de sus antiguas compañeras binarias, que se encuentran en órbitas profundas del pozo gravitacional del agujero. Además de buscar estrellas supercéleres, los astrónomos podrían buscar otras en órbita cerca del centro galáctico a velocidades que superen los 10.000 kilómetros por segundo. Velocidades orbitales tan elevadas romperían cualquier sistema estelar múltiple y normal.

Los datos que hasta ahora se han recogido en favor de un agujero negro en el centro de la Vía Láctea son bastante persuasivos. ¿Fue nuestra galaxia un cuásar? Casi seguro que no: un agujero negro de menos de tres millones de masas solares nunca podría haber producido la energía necesaria. Pero alguna actividad parecida, menos energética, sí formó parte, probablemente, de la historia de nuestra galaxia. Por tanto, el estudio de la Vía Láctea y sus vecinos cercanos abrirá nuevas perspectivas sobre las fases precoces de las galaxias y sobre la naturaleza de algunos de los objetos más potentes y exóticos del universo.

BIBLIOGRAFIA COMPLEMENTARIA

EVIDENCE FOR A SUPERMASSIVE BLACK HOLE IN THE NUCLEUS OF M31. J. Kormendy en *Astrophysical Journal*, vol. 325, n.º 1, págs. 128-141; 1988.

SUPERMASSIVE BLACK HOLES. Dirigido por Minas Kafatos. Cambridge University Press, 1988.

EL PARADIGMA DE LA MEMBRANA EN LOS AGUJEROS NEGROS. Kip S. Thorne y Richard H. Price en *Investigación y Ciencia*, junio de 1988, pág. 32.

THE RISE AND FALL OF QUASARS. M. Schmidt en *Highlights of Astronomy*, vol. 8. Kluwer Academic Publisher, 1989.

CONSTRAINTS ON THE MASS DISTRIBUTION NEAR THE CENTER OF M31 AND M32. Douglas Richstone, Gary Bower y Alan Dressler en *Astrophysical Journal*, vol. 353, n.º 1, parte 1, págs. 118-122; 10 de abril de 1990.

"DEAD QUASARS" IN NEARBY GALAXIES? Martin J. Rees en *Science*, vol. 247, n.º 4944, págs. 817-823; 16 de febrero de 1990.

Terapia génica

Se está preparando el terreno para acometer el tratamiento de las enfermedades hereditarias mediante la introducción de genes sanos. La terapia será eficaz si se logra la permanencia de los genes insertos

Inder M. Verma

Uno de cada 100 niños nace con algún defecto genético grave. Los daños suelen manifestarse en la infancia y, con mucha frecuencia, producen taras físicas o mentales, dolores y muerte prematura. Se conocen más de 4000 trastornos hereditarios, que, en su inmensa mayoría, carecen de una terapia del todo eficaz.

No debe sorprender que los científicos sueñen desde hace tiempo con introducir genes sanos en los pacientes para curar enfermedades hereditarias. Los avances experimentados en las técnicas de recombinación del ADN, que han posibilitado el aislamiento de muchos genes, y los nuevos conocimientos sobre regulación génica inducen a suponer cercano el día en que ese sueño se haga realidad.

En septiembre del año pasado dio comienzo, en los Estados Unidos, el primer ensayo clínico autorizado de terapia génica para una enfermedad hereditaria. R. Michael Blaese, W. French Anderson y sus colaboradores, del Instituto Nacional de la Salud (NIH, de sus siglas en inglés), se aprestan a introducir el gen responsable de la enzima desaminasa de la adenosina (DAA) en niños afectados por inmunodeficiencia combinada grave (SCID). La alteración de ese gen debilita el sistema inmune y es responsable de casi el 25 por ciento de los casos de esa enfermedad rara.

El plan emprendido por el grupo del NIH requiere la repetición de los tratamientos a lo largo de toda la vida. No es, por tanto, ninguna curación. A pesar de ello, el ensayo podría representar el inicio de una nueva era de la medicina. El curso de las investigaciones permite aventurar que, el siglo que viene, se pondrán en marcha ensayos clínicos de terapias génicas para un amplio abanico de enfermedades, hereditarias o no.

INDER M. VERMA enseña biología molecular y virología en el Instituto Salk de La Jolla y en la Universidad de California en San Diego. Recibió el doctorado en bioquímica por el Instituto Weizmann de Ciencias y completó su formación en el de Tecnología de Massachusetts.

Los genes pueden transferirse a las células germinales (espermatozoides, óvulos o embriones tempranos) o a las somáticas (células cuyo destino no es ser óvulo ni espermatozoide). Sin embargo, la terapia basada en la línea germinal no se contempla como opción en un futuro previsible, debido en parte a que los genes se transmitirían de generación en generación, lo que plantea profundos problemas de naturaleza ética.

Por citar algunos: ¿Se aplicaría la terapia para mejorar la raza, y no sólo para prevenir una enfermedad hereditaria? ¿A quién correspondería decidir sobre el particular? ¿Está dispuesta la sociedad a afrontar el riesgo que supone introducir cambios en el acervo génico, que podrían, en última instancia, ser deletéreos? ¿Tenemos derecho a inmiscuirnos en la evolución humana? La perspectiva de una terapia génica de las células somáticas es menos problemática, por la razón principal de que sólo afecta al propio paciente.

Los candidatos más prometedores para la terapia de células somáticas son los trastornos causados por el deterioro de un gen que haya sido aislado y clonado y que, por tanto, esté disponible para el trasplante. Esas enfermedades serían corregibles con más facilidad que las producidas por múltiples genes, o las que resultan de una alteración global, como es la pérdida o adición de un cromosoma entero. (Las células humanas portan una serie de 23 cromosomas heredados de la madre junto con la correspondiente serie heredada del padre. Cada cromosoma es un largo segmento de ADN que aloja miles de genes.)

En un mundo ideal, las enfermedades curarían de por vida, mediante un solo tratamiento y sin efectos secundarios. Y la inserción de un gen, en un cromosoma de una célula somática deseada, se llevaría a cabo en su sitio específico: mediante un proceso de recombinación homóloga, el gen sano o "terapéutico" reemplazaría, con precisión, la copia dañada. La inserción exacta aumenta la probabilidad de que un gen terapéutico funcione correctamente. Reduce también la posibilidad de que una inserción aleatoria active un oncogén (inductor de cáncer) hasta entonces quiescente o inactive a un gen supresor de cáncer.

La dura realidad, sin embargo, es que se hace dificilísimo controlar la suerte que correrá el ADN introducido en las células. Por cada gen acoplado en su lugar correcto, más de 1000 se integran aleatoriamente en el genoma (término que designa el ADN total de una célula). Los trabajos de Mario Capecchi, de la Universidad de Utah, sugieren que los obstáculos para enviar el gen a su sitio correcto, aunque grandes, no son insalvables. Mientras tanto, muchos laboratorios, entre ellos el mío en el Instituto Salk, de La Jolla, trabajan en el desarrollo de una terapia basada en el refuerzo de la función; en cuya virtud, un gen sano sustituye al producto de un gen defectuoso o inexistente, sin reemplazar físicamente el propio ADN afectado.

El refuerzo de la función puede resultar beneficioso cuando un desarreglo genético provoca una escasa inducción de proteína, si alguna. (Cada gen determina o porta instrucciones para una proteína.) Se produce un bajo rendimiento cuando, por mutación, se afecta la actividad de las dos copias, materna y paterna, de un gen,

o cuando se hereda un gen defectuoso en el único cromosoma X que porta el varón. (Las células masculinas portan un cromosoma X y otro Y; las de las mujeres, dos cromosomas X.)

Por otra parte, sin embargo, la terapia basada en el refuerzo de la función podría convertirse en contraproducente cuando una mutación provoca la superproducción de una proteína o la síntesis de una sustancia destructiva, como en el caso de la anemia falciforme. Para corregir estos defectos, la terapia debería conjugar la introducción de un gen sano y otro capaz de inactivar la versión mutada.

Por ahora, la estrategia de la mayoría de los científicos interesados en el aumento o refuerzo de la función consiste en extraer células de los pacientes, introducirles un gen terapéutico y devolverlas al sujeto. Algún día, sin embargo, los médicos podrán inyectar directamente genes ligados a sustancias que dirigirán esos genes hacia las células deseadas.

Para nuestra fortuna, la eficacia de la terapia no requiere corregir el defecto genético en cada una de los billones de células que hay en el cuerpo. En primer lugar, aunque todas las células somáticas de un individuo porten los mismos cromosomas, ciertos genes sólo funcionan en determinado tipo celular. El tratamiento, por tanto, se ceñirá exclusivamente a ese tipo. En segundo lugar, aun cuando un defecto congénito provoque la síntesis insuficiente de una proteína presente en la generalidad de las células, muchas de éstas compensan la pérdida. Así, un defecto en el gen de la DAA, que afecta parcialmente a la mayoría de las células somáticas, sólo es devastador para ciertos componentes del sistema inmunitario.

La introducción inespecífica de genes en las células se acomete por transfección (medios químicos o físicos) o por transducción (mediante virus). En los ensayos químicos, se mezclan copias abundantes de ADN correspondientes al gen sano con una sustancia cargada: por lo común, fosfato cálcico, DEAE-dextrano o ciertos lípidos. Se introduce luego esa mezcla en las células receptoras. Los compuestos químicos degradan la membrana celular y transportan el ADN hasta el interior.

El procedimiento, muy simple, ofrece el inconveniente de un bajo rendimiento. Lo usual es que sólo en una de cada 1000 a 100.000 células se integre el gen de interés en su genoma. Un médico habría de obtener un número imposible de células de los pacientes para garantizar la apropiada alteración de los millones requeridos para la terapia.

Debo señalar que la integración no es siempre decisión para la expresión génica (así se llama la producción de la proteína determinada por el gen). Sin embargo, un gen que se ha integrado tiene mayor probabilidad de permanecer en la célula. Además, se



1. BURBUJA ESTERIL que protege a David, un niño que en los años setenta padecía inmunodeficiencia combinada grave (SCID); es ésta una enfermedad hereditaria en la que el sistema inmune está gravemente dañado. Los pacientes que hoy la sufren cuentan con mejores perspectivas, que pueden aumentar aún más en el futuro. Los primeros ensayos clínicos de terapia génica se proponen mitigar una forma de esa misma enfermedad.

replicaría siempre que lo hiciese el resto del ADN, cuando la célula se prepara para dividirse. De esa manera, el gen terapéutico lo heredarían las células hijas, y las hijas de éstas, y así sucesivamente, asegurándose el suministro del producto durante la vida entera del paciente.

Entre los métodos físicos citaré la microinyección con una fina aguja de cristal y la electroporación (exposición de las células a un choque eléctrico). Por choque las células devienen permeables al ADN del medio circundante, aunque también puede dañarlas gravemente. La microinyección resulta muy eficaz; hasta una de cada cinco células incorpora el gen foráneo de forma permanente. Ahora bien, porque sólo se puede inyectar una célula cada vez, este método laborioso y tedioso no es el más aconsejable para fines terapéuticos.

La última estrategia saca provecho de la capacidad innata de los virus para entrar en las células, llevando con ellos su propio material genético. Se los ha manipulado en el laboratorio para convertirlos en vectores, o vehículos para la transferencia génica. Los virus se clasifican en función de su material genético: ADN o ARN. Las dos sustancias presentan importantes diferencias químicas, aunque las dos se construyen con las mismas subunidades, los nucleótidos, y las dos incluyen también señales reguladoras, además de las secuencias que determinan la síntesis de proteínas.

Muchos virus de ADN que pueden aceptar material genético foráneo tienen severas limitaciones en cuanto al número de nucleótidos que pueden transportar y a la gama de células que logran infectar. Otros virus de ADN poseen mayor capacidad, mas, por diversas razones, no son buenos vehículos. Además, los virus de ADN no suelen integrar su material genético en los cromosomas de las células infectadas.

Igual que sucede con los de ADN, la mayoría de virus de ARN tampoco sirven para la terapia génica, por la razón de que el ARN, incapaz de integrarse en el ADN de las células humanas, se degrada muy pronto. Los retrovirus son una excepción; convierten éstos su ARN en ADN en las células infectadas, y lo introducen después en un cromosoma. Una vez integrado, el ADN dirige la síntesis de proteínas víricas. Los retrovirus pueden transportar más material genético foráneo que algunos virus de ADN, y pueden infectar también a una amplia gama de especies y tipos celulares.

Por todo ello, los retrovirus son, de los sistemas para introducir genes estudiados hasta el momento, los más prometedores. Si no se indica expresamente, todos los sistemas de transferencia génica, todos los vectores, aquí marginados, recurren a los retrovirus.

Los retrovirus no se hallan exentos de inconvenientes. Por citar uno, sólo integran su ADN en cromosomas de células que se dividen activamente. Pero muchas células no se dividen; verbigracia, las neuronas maduras. Este tipo de células, por tanto, no son fáciles de manipular genéticamente con vectores retrovíricos.

Más preocupante es la posibilidad de que los retrovirus produzcan cáncer. El riesgo es extremadamente bajo si nos ceñimos a los propios vectores, pero aumenta si se facilita la multiplicación del virus en el cuerpo y su propagación de célula en célula. Uno de los principales retos a vencer ha sido el de desarrollar estrategias para impedir la reproducción de tales vectores.

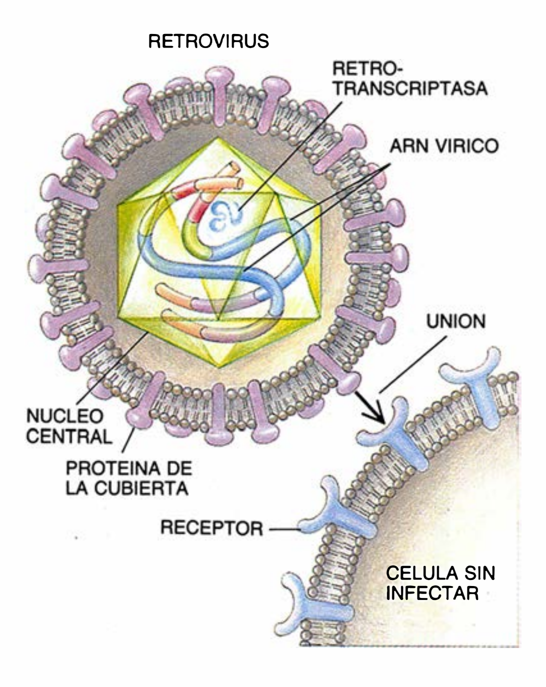
Gracias al esfuerzo de varios laboratorios se ha perfilado un sistema que parece funcionar bien [*véase la figura 3*]. Los organismos producidos con ese método presentan una cubierta externa normal, con todas las proteínas víricas. El ARN del retrovirus, sin embargo, carece de instrucciones para sintetizar tales proteínas. Los genes terapéuticos incluidos se ocupan de llevar a cabo esas instrucciones omitidas.

Utilizando su cubierta, los virus entran en las células y descargan el contenido vírico en el citoplasma. A continuación, las enzimas víricas convierten el ARN en ADN y facilitan la integración de dicho ADN en el genoma de la célula hospedante. Aquí termina su andadura el virus.

En circunstancias normales, el ADN retrovírico integrado –profago– dirigirá la síntesis de proteínas víricas y ARN, que se ensamblarán para originar un clon de virus similares al original. Los retrovirus alterados, por el contrario, privados de instrucciones para sintetizar proteínas víricas, no producen descendencia. En definitiva, los virus desaparecen de la célula, dejando como único vestigio al gen foráneo, con las secuencias nucleotídicas que regulan su expresión.

Aunque los retrovirus pueden infectar muchos tipos de células, sólo algunos de éstos cumplen los requisitos para su manipulación genética. Las células deben ser lo suficientemente fuertes como para resistir las manipulaciones, y deben poder extraerse del cuerpo y retornar a él con razonable facilidad. Además, deben tener una vida media larga, de meses o años, aunque lo preferible es que duren toda la vida del paciente. Las células de la médula ósea, piel e hígado son las que mejor cumplen esos criterios; por ello, las enfermedades que pueden tratarse mediante manipulación son las candidatas mejores para la terapia génica.

Las células de la médula ósea, hematopoyéticas, pueden, en principio, emplearse para corregir defectos causados por alteraciones genéticas de



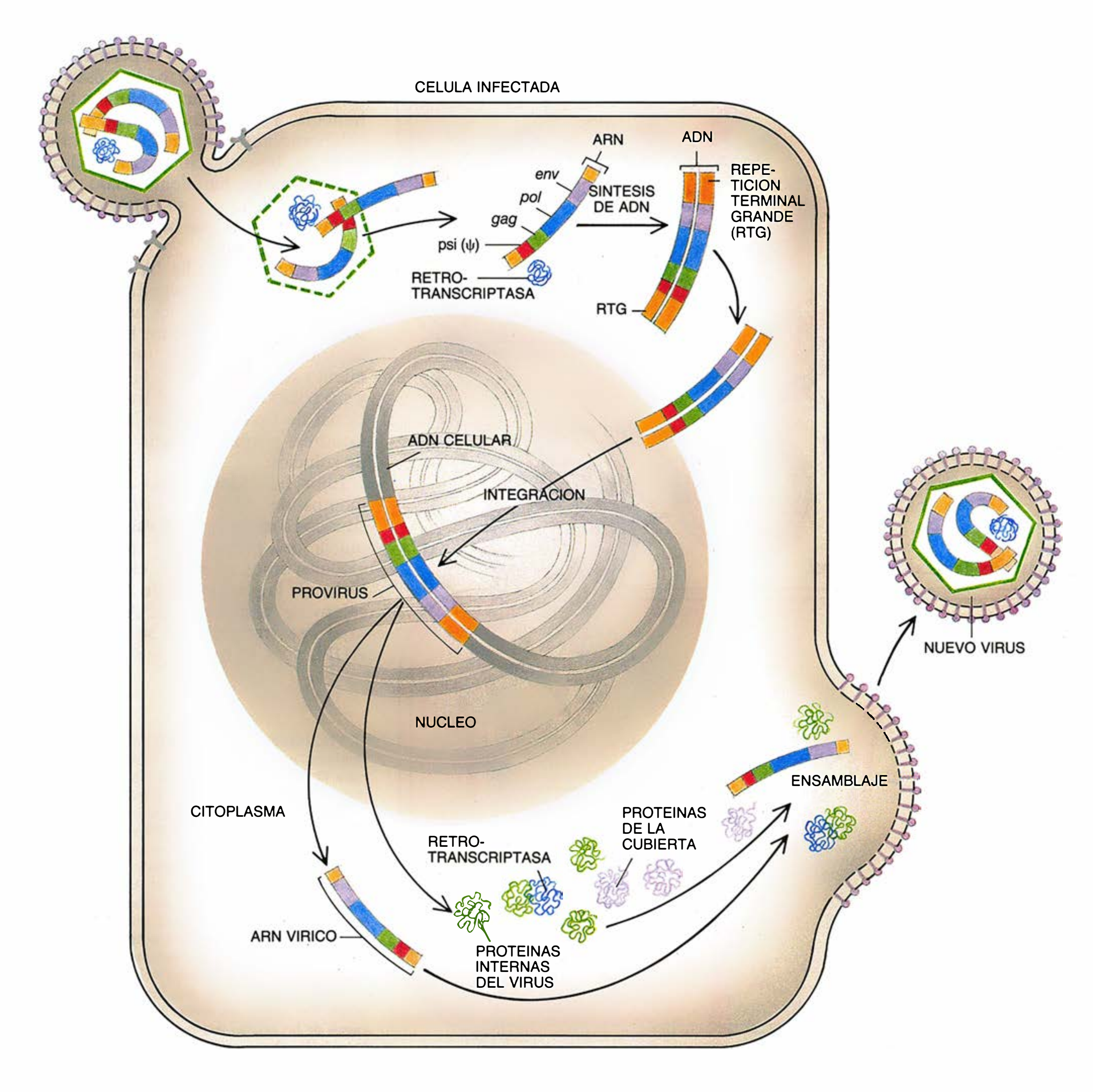
2. CICLO DE VIDA de un retrovirus. Comienza cuando el virus se adhiere a una célula *(arriba)*, entra en ella *(derecha)* y le inyecta su material genético (ARN) y proteínas en el citoplasma. El ARN retrovírico típico incluye tres regiones: *gag (verde)*, *pol (azul)* y *env (púrpura)*, que determinan, respectivamente, proteínas del interior del virus, la enzima retrotranscriptasa y los componentes de la cubierta. Posee también tres dominios sin información para proteínas: dos en los extremos *(naranja claro)* y otro llamado psi, ψ *(rojo)*. En el citoplasma, la retrotranscriptasa convierte el ARN en ADN. Los extremos de la molécula de ADN, también llamados repeticiones terminales grandes *(naranja oscuro)*, condicionan la actividad de los genes víricos y facilitan la inserción del ADN vírico en el ADN celular. El ADN así instalado (provirus) dirige la síntesis de ARN y proteínas víricas; éstas engloban al ARN, formando nuevas partículas víricas, que salen de la célula por gemación.

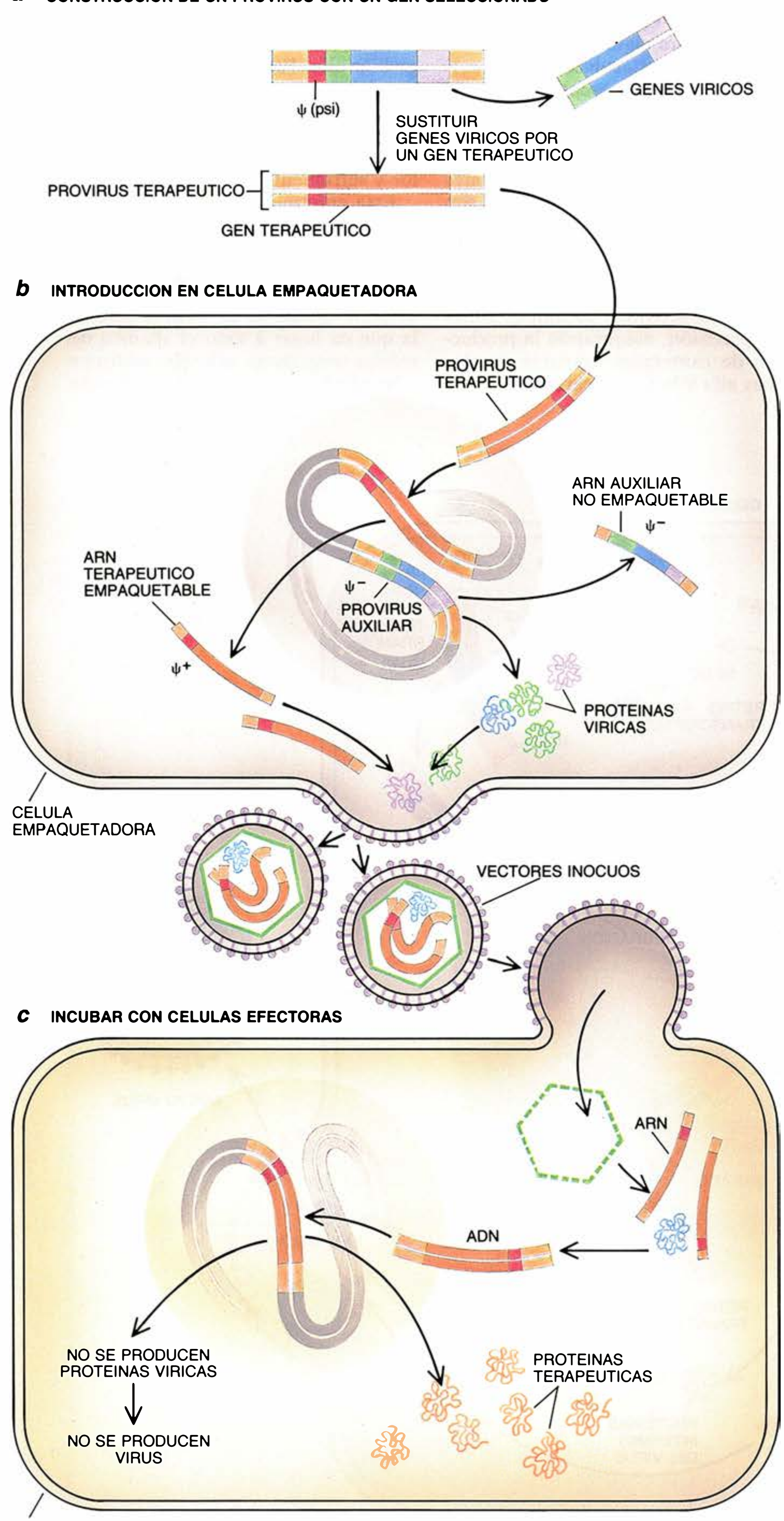
los eritrocitos o los leucocitos (importantes para la inmunidad). El SCID, causado por una deficiencia de la DAA, es una de las enfermedades hereditarias que afectan a las células inmunes; otra es la deficiencia en la adhesión de los leucocitos, circunstancia que determina una pobre movilización de los leucocitos y que favorece las infecciones recurrentes. Entre las enfermedades asociadas con una deficiencia de los glóbulos rojos están las talasemias, causadas por un desequilibrio en los genes que determinan las subunidades de la molécula de hemoglobina, el transportador de oxígeno de los eritrocitos.

La talasemia beta se convirtió en la primera enfermedad con expectativa de tratamiento con terapia génica. Su historia ilustra algunos de los problemas que han obstaculizado los esfuerzos empeñados en asentar la terapia génica y, en particular, la basada en la médula ósea.

Los glóbulos rojos de los pacientes afectados de betatalasemia son deficientes en globina beta, que en los individuos sanos se combina con hierro (hemo) y globina alfa para producir hemoglobina. Las células sanas regulan la actividad de ambos genes con precisión, asegurando la producción de cantidades similares de globinas alfa y beta. La carencia de globina beta no sólo priva de la producción de hemoglobina, sino que acarrea, además, un exceso relativo de globina alfa. Este exceso acelera, a su vez, la muerte celular y puede ocasionar anemias graves. Los afectados acaban sucumbiendo a la enfermedad antes de los 20 años, tras varios de dolor y sufrimientos.

Esta enfermedad y otras alteraciones hereditarias de la sangre podrían probablemente tratarse con eficacia introduciendo genes sanos en las células madre de la médula ósea. Esta subserie celular de la médula ósea es la que da lugar a todo el abanico de células sanguíneas; además, sustituye a las células que mueren a lo largo de





3. VECTORES RETROVIRICOS que se ensamblan, o empaquetan, en células diseñadas para liberar sólo vectores inocuos. Los investigadores sustituyen genes del provirus por otro terapéutico *(a)* e insertan el provirus en una célula empaquetadora *(b)*. El ADN vírico dirige la síntesis de ARN del virus, pero como le faltan sus propios genes, no puede producir las proteínas necesarias para empaquetar el ARN en partículas que infectarían a otras células. Las proteínas de que se carece son suministradas por un provirus "auxiliar", al que se le ha eliminado la región psi, región que es fundamental para incluir el ARN en las partículas víricas; sin psi, no se pueden formar virus que lleven el ARN auxiliar. Por tanto, las únicas partículas que escapan de las células son las que portan el ARN terapéutico. Estas partículas pueden introducirse en otras células *(c)* e insertar el gen terapéutico en el ADN celular, pero no reproducirse.

toda la vida de una persona. La introducción estable del gen deseado en una célula madre garantizaría la producción de células sanguíneas normales durante toda la vida del paciente.

Por desgracia, las células madre humanas, que abundan, oponen una resistencia invencible al aislamiento. Los investigadores se han visto, por tanto, obligados a recurrir a una estrategia menos eficaz: infectar grandes cantidades de células de la médula ósea con un retrovirus terapéutico, con la esperanza de infectar un número suficiente de células madre.

A los estudios sobre la globina beta debemos buena parte de las pruebas que demuestran la validez relativa de ese planteamiento. Varios laboratorios, apoyándose en vectores retrovíricos, han corroborado la persistencia estable, en células de médula ósea de ratón, de un gen para la globina beta humana. Y Richard C. Mulligan y sus colaboradores, del Instituto Whitehead de Investigaciones Biomédicas de Cambridge, han demostrado, además, que dicho gen se expresa cuando las células de que hablamos se implantan en ratones.

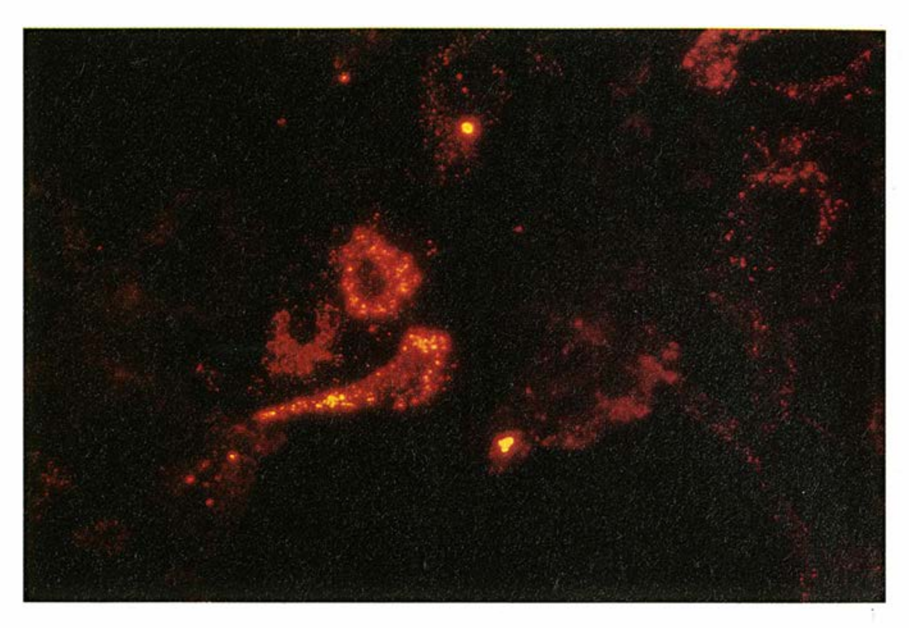
Sin embargo, nadie ha sido capaz de obtener niveles significativos de síntesis de globina en los animales receptores. Este problema ha sido un contratiempo, aunque un descubrimiento realizado por F. G. Grosveld y sus colaboradores, del Instituto Nacional de Investigaciones Médicas de Londres, permite abrigar esperanzas de solución.

Ellos han identificado unos fragmentos de ADN (intensificadores), situados a miles de nucleótidos de distancia del propio gen, que en los glóbulos rojos normales elevan notablemente la producción de ARN mensajero de la globina. El ARN mensajero se transcribe, o se copia, del ADN, y es el molde para sintetizar

proteínas. Por tanto, una cantidad elevada de ARN mensajero indica que la proteína correspondiente se está produciendo en abundancia. Parece razonable pensar que, si se colocan juntos los intensificadores específicos de la globina y el gen de globina en un vector retrovírico, mejoraría la síntesis de globina en el cuerpo. Actualmente se están realizando experimentos para comprobarlo.

En general, las células de médula ósea genéticamente alteradas también han producido, *in vivo,* una expresión pobre de otros genes. El problema exige resolverse si queremos ver hecha realidad la terapia génica basada en las células de médula ósea.

Unido a un nivel aceptable de expresión génica, sería de desear una actividad duradera. Los descubrimientos recientes relacionados con la globina apuntan la posibilidad de que cueste menos lograr una prolongada expresión de genes insertados en la médula ósea que alcanzar altos niveles de síntesis de proteínas. En ese contexto, Chung L. Li y V. J. Dwarki, en mi laboratorio, han conseguido una expresión continua, aunque débil, del gen de la globina beta humana en ratones, durante al menos un período de estudio de cinco meses, equivalente a 15 o 20 años en un ser humano. El gen de la globina alfa se mantiene funcional durante al menos 10 meses.



4. CELULAS HEPATICAS de conejos genéticamente deficientes en el receptor de la lipoproteína de baja densidad (LDL), expresando dicho receptor *(regiones brillantes)* después de recibir desde fuera su gen. Este resultado positivo permite pensar en la posibilidad de que un defecto genético humano similar, que provoca una acumulación de colesterol, sea tratado con terapia génica. James M. Wilson, del Instituto Médico Howard Hughes de la Universidad de Michigan en Ann Harbor, y J. Roy-Chowdhury, de la Facultad de Medicina Albert Einstein, hicieron la fotomicrografía.

Otros descubrimientos, surgidos a partir del trabajo sobre la talasemia beta, resaltan la complejidad añadida cuando la corrección de una enfermedad depende de que la expresión del gen terapéutico esté regulada con exactitud. Para muchas enfermedades, incluida la SCID, la mera producción de algo de proteína antes inexistente es mejor que nada. Pero eso no vale para la talasemia. Como un exceso relativo de globina alfa o beta puede dañar a las células, la actividad del gen de globina terapéutico debe ser idéntica a la de una versión normal. Para nuestra desgracia, queda mucho por conocer de los mecanismos que controlan la actividad de la mayoría de los genes, y entre ellos el de la beta globina. Pero no se detiene el avance de cuanto ayuda a mejorar el diseño de los vectores para la terapia génica.

Los investigadores del NIH que trabajan en la SCID han abordado la terapia génica basada en las células de médula ósea dando un rodeo, en parte para eludir el problema de la menguada expresión. Durante los ensayos, los pacientes son tratados con una subserie de linfocitos *T* circulantes, leucocitos que desempeñan un papel crucial en el sistema inmune. Las células *T* son devastadas cuando falta la enzima desaminasa de adenosina.

Los linfocitos alterados mediante retrovirus son introducidos en niños que reciben una ayuda adicional con inyecciones de DAA-PEG-DAA mezclada con polietilenglicol, para incrementar la vida media de la enzima. El éxito de este tratamiento se evaluará comparando las mejoras obtenidas en la función inmune con las alcanzadas con el mero reemplazamiento enzimático. Lamentablemente, las células *T* no poseen la longevidad de las células madre, por lo que la enfermedad no puede curarse de manera indefinida con un solo tratamiento.

La existencia de tratamientos no genéticos para el SCID (incluido el trasplante de médula ósea) plantea la cuestión general de si está justificado el uso de una terapia génica experimental con pacientes, cuando existen otras alternativas. De acuerdo con la opinión mayoritaria, dicha experimentación es aceptable si se demuestra que los riesgos son bajos y si, por otro lado, la terapia génica promete ser significativamente más eficaz que los tratamientos existentes, o si al paciente no se le pueden aplicar los tratamientos establecidos. En el caso de la SCID, por ejemplo, no todos los pacientes tienen acceso a médula ósea de un donante compatible.

La alteración genética de los linfocitos o de las células de médula ósea se realiza para corregir defectos en esas mismas células o en sus descendientes. Se estudian también células de la piel como un objetivo bien distinto: la síntesis y secreción de proteínas que suelen sintetizarse en un tipo celular, y que son transportadas por el plasma sanguíneo para ser utilizadas por otras.

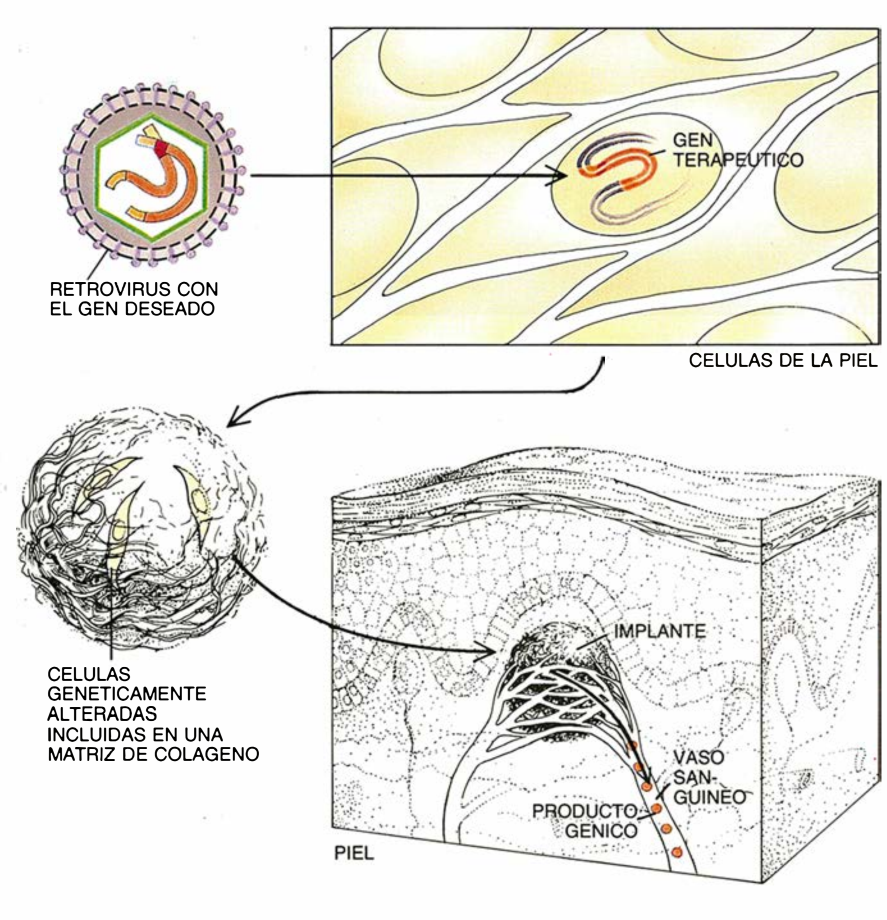
En principio, la implantación de células de la piel podría corregir muchos defectos. Por ejemplo, la hemofilia (causada por una falta de factores coagulantes de la sangre producidos por el hígado) y enfermedades que resultan de la síntesis insuficiente de ciertas hormonas (la del crecimiento por ejemplo). Determinadas enfermedades que tienen su origen en la deficiente producción de proteínas que se sintetizan en muchos sitios constituirán, a su vez, buenas candidatas si lo tejidos más afectados por la deficiencia pudiesen tomar las proteínas que necesitan de la sangre.

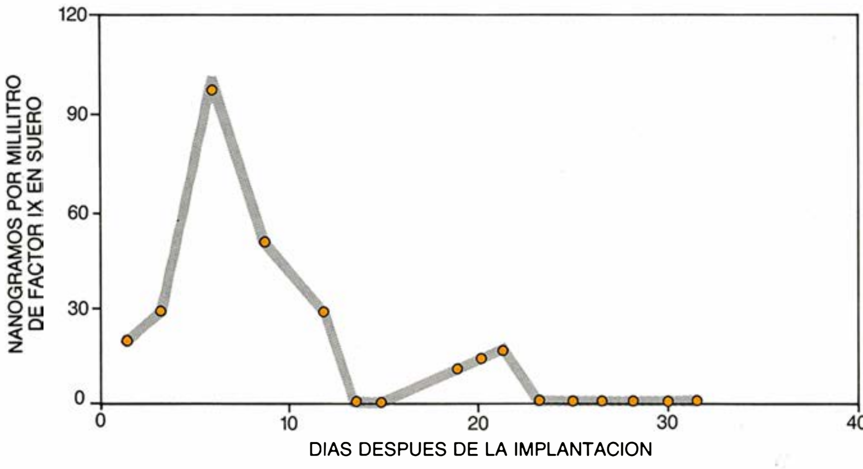
Los fibroblastos, un componente de la dermis (capa inferior de la piel), son los más adecuados para la terapia, que en este caso supondría la implantación de las células alteradas nuevamente en la dermis. Asequibles y fuertes, se multiplican fácilmente en el laboratorio. Además, pueden secretar sustancias en la sangre, y serían fáciles de extirpar llegado el caso.

Mi laboratorio ha estudiado exhaustivamente la utilidad de los fibroblastos de la piel en el tratamiento de

la forma de hemofilia causada por la falta del factor de coagulación IX, producido por el hígado. Nuestros resultados subrayan el gran potencial terapéutico de dichas células.

En uno de nuestros estudios, A. Dusty Miller, ahora en el Centro de Investigaciones Oncológicas Fred Hutchison, en Seattle, en colaboración con George C. Brownlee y Don S. Anson, de la Universidad de Oxford, demostraron que podía obligarse a los fibroblastos a sintetizar y secretar factor IX, aun cuando ellos no fabrican dicha proteína. (No sabemos si ocurrirá lo mismo con otras proteínas.) Además, Daniel C. St. Louis, Jonathan H. Axelrod y Raphael Scharfmann, de mi grupo, utilizaron retrovirus para insertar el gen del factor IX humano en fibroblastos; implantaron posteriormente las células en la dermis de ratones, desde donde comenzaron a liberar el factor en la sangre.





5. CELULAS DE LA PIEL, en las que se ha introducido un gen foráneo, se incluyen en una matriz de colágeno y posteriormente se implantan en la dermis para que liberen el producto génico en la sangre *(arriba)*. En un experimento preliminar, desarrollado en el laboratorio del autor en La Jolla, se implantaron en ratones fibroblastos de piel con el gen humano para el factor IX, una proteína que participa en la coagulación de la sangre, y que normalmente es secretada por el hígado. La vascularización fue buena; los fibroblastos secretaron el factor humano durante unas dos semanas, aproximadamente *(gráfico)*. En la actualidad, la eficacia ha mejorado bastante, y se consigue que los implantes de fibroblastos liberen proteínas foráneas durante períodos más largos de tiempo.

Este estudio no sólo demostró que era posible la expresión del factor IX en animales, sino que además nos enseñó una importante lección. Aproximadamente 15 días después de implantar las células, el factor humano desapareció de la sangre de los ratones. Parecía que los animales receptores habían desarrollado una respuesta inmune contra la proteína humana foránea. Moraleja: la terapia génica funcionará mejor, probablemente, en pacientes que sean capaces de sintetizar, al menos, una pequeña cantidad de la proteína deficiente. Si no es así, el sistema inmune podría rebelarse contra el producto de un gen insertado.

Encontramos indicios que daban pie a suponer que, a diferencia de las células de médula ósea estudiadas hasta ahora, los fibroblastos podrían producir suficiente cantidad del producto seleccionado para corregir una enfermedad. La extrapolación de los datos obtenidos con ratones permite pensar que un implante del tamaño de una moneda de cinco duros puede producir proteína en cuantía suficiente para aliviar una deficiencia de factor IX en el hombre. En colaboración con Kenneth M. Brinkhows, de la Universidad de Carolina del Norte en Chapel Hill, vamos a estudiar la posibilidad de corregir la hemofilia en perros mediante implantes de fibroblastos. Si tuviéramos éxito, estarían justificados los ensayos en humanos.

Los fibroblastos genéticamente alterados podrían implantarse también en el cerebro, para corregir trastornos nerviosos. No es fácil abordar el cerebro: muchas drogas que circulan por la sangre no llegan hasta él. Además, no podemos extraer neuronas para modificarlas genéticamente, sin que se resienta el cerebro. Los fibroblastos manipulados podrían utilizarse para que, en teoría, secretasen proteínas que, por difusión, llegaran a las células nerviosas.

Los primeros resultados son alentadores. Fred Gage, de San Diego, ha demostrado que los implantes manipulados para secretar el factor de crecimiento nervioso estimulaban el crecimiento neuronal en cerebros de rata. La regeneración se producía en

ENFERMEDAD	INCIDENCIA	PRODUCTO NORMAL DEL GEN DEFECTUOSO	CELULAS EFECTORAS	SITUACION ACTUAL
Hemoglobinopatías (talasemias)	1 de cada 600 en ciertos grupos étnicos	Componentes de la hemoglobina	Células de médula ósea (origen de la sangre)	La producción de globina en los animales receptores debe mejorarse
Inmunodeficiencia combinada grave (SCID)	Rara	Desaminasa de la adenosina (DAA) en aproximadamente una cuarta parte de los pacientes con SCID	Células de médula ósea o linfocitos *T*	Los ensayos clínicos de la terapia linfocítica para la deficiencia de DAA están en marcha
Hemofilia A, Hemofilia B	1 de cada 10.000 varones; 1 de cada 30.000 varones	Factor coagulante VIII; Factor coagulante IX	Células hepáticas o fibroblastos	Buenas posibilidades de ensayos clínicos (con fibroblastos) en los próximos cinco años
Hipercolesterolemia familiar	1 de cada 500	Receptor hepático de la lipoproteína de baja densidad (LDL)	Células hepáticas	Los estudios con animales están en sus etapas preliminares
Enfisema hereditario	1 de cada 3500	Antitripsina-alfa$_1$ (producto hepático que protege a los pulmones de degradaciones enzimáticas)	Células hepáticas o pulmonares	Trabajos muy preliminares
Fibrosis cística	1 de cada 2500 caucasianos	Sustancia importante para mantener los alveolos pulmonares libres de mucosidades	Células pulmonares	Inhalaciones directas del gen a los pulmones, como posibilidad teórica
Distrofia muscular de Duchenne	1 de cada 10.000 varones	Distrofina (componente estructural del músculo)	Células musculares (en particular las embrionarias que se desarrollan como fibras musculares)	Trabajos preliminares. Genes no distróficos inyectados en músculo dirigen la síntesis de la proteína correspondiente
Enfermedades relacionadas con el almacenamiento lisosómico	1 de cada 1500 presenta alguna forma de la enfermedad	Enzimas que degradan moléculas complejas en los compartimentos intracelulares que reciben el nombre de lisosomas	Varía, según la enfermedad	La mayoría de las enfermedades necesitarán el envío del gen a células del cerebro (una tarea complicada), así como a otros tipos celulares

6. CANDIDATAS POTENCIALES para los primeros intentos de terapia génica serían las enfermedades causadas por defectos en un solo gen, que además esté ya clonado. En pocas palabras, los médicos deben extraer células del paciente, insertar en ellas un gen sano y devolverlas al cuerpo.

el tipo de neuronas cuya degeneración está asociada con la pérdida de memoria en la enfermedad de Alzheimer, si bien no se ha establecido el papel del factor en esa enfermedad. De la misma manera, se estudian implantes que fabrican levodopa (L-dopa), un precursor del neurotransmisor dopamina, con animales modelo de la enfermedad de Parkinson. Nadie sabe a ciencia cierta cuál es la causa del Parkinson, aunque parece intervenir la escasez de dopamina. Se investiga aún la supervivencia de los implantes de fibroblastos en la piel o en el cerebro.

Comparadas con las células de la médula ósea y la piel, las células hepáticas son unas recién llegadas al campo de la terapia génica. Pueden ser importantes en el tratamiento de enfermedades genéticas causadas por una disfunción de las células hepáticas. Mulligan, del Instituto Whitehead, y James M. Wilson, a la sazón en el mismo Instituto, y, por separado, Theodore Friedmann y sus colaboradores, en San Diego, han introducido con éxito el gen del receptor de la lipoproteína de baja densidad (LDL) en células hepáticas, y han conseguido que exprese, en el laboratorio, receptores biológicamente activos. Las células proceden de conejos Watanabe, deficitarios en el receptor de la LDL por condición genética alterada, igual que los seres humanos afectados de hipercolesterolemia familiar, una condición que puede provocar ataques al corazón.

Se ha estudiado la inyección directa en conejos Watanabe vivos de complejos formados por el gen receptor y una proteína que lo introduzca en el hígado. (La inyección directa en seres humanos evitaría la cirugía para extirpar células hepáticas.) La proteína determinada por dicho gen pudo ser detectada en el cuerpo pero, lo mismo que en los cultivos de células *in vitro*, sólo durante cierto tiempo. La longevidad podría mejorarse; las investigaciones con células hepáticas están todavía en pañales.

Aunque las células de médula ósea, hígado y piel son las que reciben la mayor atención, no son las únicas. Por ejemplo, los retrovirus pueden transportar genes de productos exportables hasta las células endoteliales, que revisten a las arterias. Esas células tienen un contacto más íntimo con la sangre que los fibroblastos; podrían enviar, por tanto, el producto con más rapidez.

Los investigadores consideran también la posibilidad de inyectar genes sanos para la distrofina (componente estructural del músculo) directamente en músculo de ratones con un defecto similar a la distrofia muscular de Duchenne. Hay razones para esperar que los genes se expresen. Otros genes inyectados en músculos de animales vivos han producido proteínas durante meses, incluso cuando el ADN no estaba integrado en los cromosomas. Podría tratarse también la fibrosis quística, una enfermedad hereditaria que afecta a los pulmones, empaquetando genes sanos en retrovirus para su inhalación a través de un atomizador.



La terapia génica no debe limitarse a reparar los efectos de genes trastornados. Puede aportar también nuevas propiedades a las células, para aumentar su potencial defensivo contra las enfermedades.

A este propósito, Steven A. Rosenberg y sus colaboradores, del Instituto Nacional del Cáncer, han demostrado que los linfocitos obtenidos de tumores de pacientes, cultivados con interleucina-2 (un activador de las células *T*), pueden detener algunos cánceres. Ahora esperan aumentar el poder anticancerígeno de esos linfocitos que se infiltran en los tumores (LIT), insertándoles un gen que determina el factor de la necrosis tumoral, una potente molécula del sistema inmune. El factor, que tiene actividad anticancerosa, no se fabrica normalmente en las células *T*. Se espera que pronto comiencen los ensayos clínicos [véase "Inmunoterapia del cáncer", por Steven A. Rosenberg; INVESTIGACIÓN Y CIENCIA, julio de 1990].

Otro grupo de investigadores intenta inducir en varios tipos celulares la producción de CD4, una molécula de las células *T* mermadas por el virus del SIDA. Para entrar en la célula, una proteína de la cubierta del virus debe interaccionar con la molécula CD4. La presencia masiva de moléculas CD4 en la sangre podría atraer a los virus y evitar que éstos entren en contacto con las células. Muchas son las ideas que se están aportando a la terapia génica, como la de inducir a las células endoteliales a secretar factores que impidan que se formen coágulos de sangre en las arterias de los pacientes después de una operación de corazón.

La idea de introducir genes para corregir enfermedades hereditarias y otras disfunciones es poco menos que revolucionaria. Quizás ésa es una de las razones por las que el progreso es más lento de lo que se esperaba en un principio. Las criaturas modernas son producto de millones de años de evolución. No cabía esperar que las primeras tentativas de insertar genes en células produjesen, de forma fácil, una expresión estable.

Para sanar enfermedades, los investigadores han de estar seguros de que los genes terapéuticos se expresan bien y se asientan de forma estable en el cuerpo. Continuamente se producen descubrimientos, como la importancia de incluir secuencias reforzantes con algunos genes en vectores retrovíricos, que indican el camino a seguir. También se necesitan mejores métodos para devolver al cuerpo las células alteradas (como las hepáticas), procedimientos para prolongar la vida de las células implantadas y técnicas para aislar células madre humanas (en vez de las actuales de médula ósea).

Al mismo tiempo, debe confirmarse la seguridad de los vectores retrovíricos, realizando para ello los estudios necesarios en animales pequeños y grandes, y deben continuar los esfuerzos encaminados a introducir nuevos elementos de protección. A pesar de que ya se utilizan retrovirus que han perdido su capacidad de replicación, existe aún cierta posibilidad de que causen cáncer. También sería conveniente que se desarrollasen alternativas a los retrovirus, y que se profundice en las investigaciones tendentes a conseguir una inserción de los genes en sus lugares específicos.

La posibilidad de sanar definitivamente enfermedades genéticas, con un único y seguro tratamiento, justifica todos los esfuerzos que se están dedicando a ello, pero debo acabar recordando que la terapia génica no puede corregir todas las enfermedades humanas. La mayoría de las aflicciones humanas no son genéticas. Son ambientales, causadas por infecciones microbianas, que se propagan debido a las deficientes condiciones sanitarias, contaminación de aguas, malnutrición y otros factores que escapan a la ingeniería genética. También esas enfermedades merecen una especial atención.

BIBLIOGRAFIA COMPLEMENTARIA

LINEAGE-SPECIFIC EXPRESSION OF A HUMAN β-GLOBIN GENE IN MURINE BONE MARROW TRANSPLANT RECIPIENTS RECONSTITUTED WITH RETROVIRUS-TRANSDUCED STEM CELLS. Elaine A. Dzierzak, Thalia Papayannopoulou y Richard C. Mulligan en *Nature*, vol. 331, n.º 6151, págs. 35-41; 7 de enero de 1988.

DISRUPTION OF THE PROTO-ONCOGENE *INT*-2 IN MOUSE EMBRYO-DERIVED STEM CELLS: A GENERAL STRATEGY FOR TARGETING MUTATIONS TO NON-SELECTABLE GENES. Suzanne L. Mansour, Kirk R. Thomas y Mario R. Capecchi en *Nature*, vol. 336, n.º 6197, págs. 348-352; 24 de noviembre de 1988.

AN ALTERNATIVE APPROACH TO SOMATIC CELL GENE THERAPY. D. St. Louis e I. M. Verma en *Proceedings of the National Academy of Sciences*, vol. 85, págs. 3150-3153; 1988.

HUMAN GENE THERAPY. Eve K. Nichols y la Dirección del Instituto de Medicina de la Academia Nacional de Ciencias. Harvard University Press, 1988.

PROGRESS TOWARD HUMAN GENE THERAPY. Theodore Friedmann en *Science*, vol. 244, n.º 4910, págs. 1275-1281; 16 de junio de 1989.

Ciencia y sociedad

Desigualdad de los idénticos

John y James Brick son médicos y viven en Virginia. También son gemelos idénticos, según se ha corroborado con una prueba de análisis bioquímico. Pero John ha tenido siempre un apetito y una cintura mayores que los de James. Ambos hermanos expusieron recientemente su caso en *New England Journal of Medicine*, revista que ha venido publicando numerosos artículos defendiendo que el peso corporal, como otros muchos rasgos humanos, está determinado por los genes. Si tal fuera cierto, se preguntaban John y James, ¿por qué uno de ellos es más gordo que el otro?

En efecto, ¿por qué? En los gemelos idénticos se fundan las investigaciones sobre la relación entre naturaleza y crianza. Parten del supuesto de que nacen iguales y atribuyen las divergencias fisiológicas y de comportamiento a la acción del medio.

Ahora bien, los gemelos idénticos pueden salir del vientre materno con notables diferencias fisiológicas e incluso genéticas. Este fenómeno, que explica quizás el caso de los Brick, no es "apreciado en absoluto" ni siquiera por los investigadores de gemelos, afirma E. Fuller Torrey, psiquiatra que ha realizado estudios sobre mellizos en el Instituto Nacional de Salud Mental y que en congresos recientes ha planteado la cuestión de la disparidad biológica.

Los gemelos idénticos, monocigóticos, surgen cuando un cigoto, u óvulo fecundado, se divide, por razones desconocidas, en dos cigotos distintos inmediatamente después de la concepción. (Los mellizos proceden de dos óvulos fecundados por dos espermatozoides.)

Torrey señala que entre el 10 y el 20 por ciento de los gemelos monocigóticos se hallan expuestos a la condición de transfusión fetal, por la que un gemelo recibe más oxígeno y nutrientes que el otro; al nacer, uno de ellos puede llegar a pesar la mitad que el otro. Además, es posible que sólo uno se vea afectado, dentro del útero, por virus, bacterias o drogas consumidas por la madre. Los efectos pueden ser permanentes. Existen casos en los que sólo un gemelo tiene el síndrome fetal del alcohol o sufre las consecuencias de la talidomida.

La disparidad genética entre gemelos monocigóticos ocurre de diversas maneras, según Torrey. En casos extremos, un mellizo pierde un cromosoma entero o recibe, por contra, cromosomas suplementarios, durante la división inicial del óvulo, o poco después. La mutación cromosómica puede provocar el síndrome de Down y otros tipos de retrasos físicos y mentales en el gemelo; puede alterar incluso su sexo.

La divergencia cromosómica, rara, se detecta con facilidad. Pero los gemelos en apariencia idénticos pueden presentar diferencias genéticas sutiles que incidan en el comportamiento o en la fisiología. Disparidad genética que no resultará tan detectable en las pruebas sanguíneas que se acometen para confirmar el monocigotismo.

La defensa cruel de las acacias

Algunos pastos no son lugares seguros para los antílopes. En muchos ranchos de ganado salvaje de Sudáfrica, cientos de kudus, un tipo de antílope, morían durante la estación seca. La causa no estaba en los virus ni en las balas. Los asesinos, según Wouter van Hoven, zoólogo de la Universidad de Pretoria, eran las hojas de acacia.

De ordinario, la acacia es un recurso alimentario básico para el kudu. Como otras muchas plantas, produce taninos para defenderse de los herbívoros. En la mayoría de los casos, los niveles de taninos no dañan al kudu. Pero en condiciones de gran presión ambiental, como ramoneo masivo o sequía, las hojas aumentan su producción de esta toxina astringente: desarrollan entonces un gusto desagradable.

Los kudus que viven en libertad buscan las hojas con niveles bajos de taninos, pero a los cercados no les queda más remedio que comer también las hojas envenenadas. Van Hoven encontró que durante la sequía las acacias de algunos ranchos producían taninos en cuantía suficiente para inactivar las enzimas hepáticas del kudu, lo que mataba a la mayoría de los animales en dos semanas.

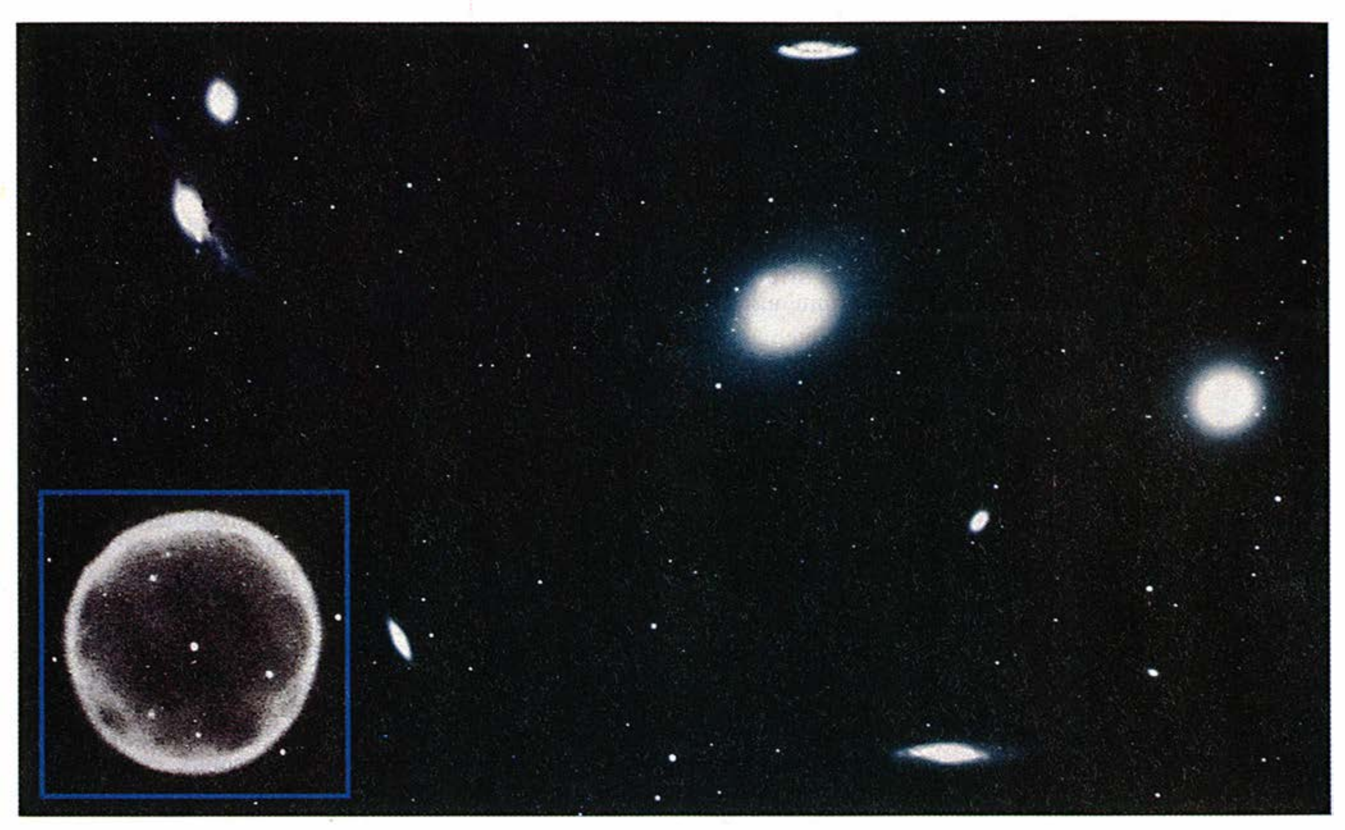
Aparentemente, las acacias no esperan a ser atacadas para movilizar sus defensas. Parecen dotadas para conspirar contra los ramoneadores mediante algún tipo de comunicación entre sí. En 1983, Jack C. Schultz e Ian Balwin, dos biólogos que por aquel entonces se hallaban en el Dartmouth College, ofrecieron las primeras pruebas de comunicación vegetal. Encontraron que los arces de azúcar sanos tendían a producir mayores cantidades de taninos y otros compuestos defensivos en presencia de plantas cuyas hojas estaban dañadas.

La investigación preliminar de van Hoven indica que, cuando un antílope ramonea en una acacia, las hojas emiten etileno. Este compuesto parece señalar a otros árboles, situados a una distancia de hasta 50 metros a favor del viento, la presencia de un herbívoro. Los árboles intactos aumentan entonces su propia producción de taninos antes de que su follaje haya sufrido daño alguno. En condiciones de laboratorio, van Hoven encontró que las hojas dañadas liberaban 20 veces más etileno que las hojas sin alterar. "Cuando expuse una planta intacta a [niveles elevados] de etileno, aumentó la concentración de taninos a los 30 minutos", observa el zoólogo de Pretoria.

El incremento de la producción de taninos a corto plazo actúa "como un mecanismo natural y regulador de la población", añade. Cuanto mayor es la frecuencia con que un árbol es ramoneado, más taninos produce. Para



1. Kudus ramoneando en un bosquecillo de acacias. Foto: Animals Animals.



2. Galaxias del cúmulo de Virgo. Se hallan sólo a 48 millones de años luz. A esa cifra se ha llegado después de buscar nebulosas planetarias en estas galaxias y en antiguas estrellas rodeadas de nubes esféricas de gas.

solucionar el problema, los rancheros redujeron el número de kudus en sus pastos, alimentándolos con alfalfa cuando la sequía aprieta.

Medir la eternidad

¿Llegarán algún día a ponerse de acuerdo los astrónomos sobre la edad del universo? Durante años se han barajado cifras muy dispares sobre la misma, con una banda situada entre 10.000 y 20.000 millones de años. Las mediciones recientes han convencido a muchos astrónomos, aunque no a todos, de que el cosmos no ha cumplido los 12.000 millones de años.

De confirmarse, estos datos obligarían a revisar las teorías dominantes sobre la naturaleza del universo. "Un universo joven crea más problemas que resuelve", afirma George H. Jacoby, de los Observatorios Nacionales de Astronomía Optica y partícipe en estas nuevas medidas.

Edwin P. Hubble fue el primer astrónomo de la época moderna en atribuirle al universo una edad finita. Hace unos 60 años descubrió que la luz de las galaxias tendía a derivar hacia el rojo, señal de que se alejan de nosotros, en proporción directa a su distancia relativa respecto de la Tierra. Una galaxia situada a doble camino que otra se aleja con doblada celeridad respecto a su homóloga. Esta relación entre distancia y velocidad de recesión llevó a Hubble a postular la expansión del universo.

La tasa de expansión del cosmos se expresa mediante la constante de Hubble. Se calcula dividiendo la velocidad de recesión de una galaxia por su distancia a la Tierra. Para estimar la edad del universo, basta con invertir la constante de Hubble, y dividir, pues, la distancia por la velocidad de recesión. (Este método presupone que el universo se ha expandido siempre al mismo ritmo desde la gran explosión hasta hoy.)

Las velocidades de recesión se calculan, sin mayor dificultad, a partir de su corrimiento hacia el rojo. Sí cuesta, por contra, medir la distancia real de las galaxias, la más cercana de las cuales se halla a millones de años luz de la Tierra. Además, el corrimiento hacia el rojo de cualquier galaxia y, en particular, de las más próximas, puede verse distorsionado por el "movimiento peculiar" que resulta de la atracción gravitatoria de otras galaxias. Para ajustar con precisión el valor de la constante de Hubble y la edad del universo, deben obtenerse las distancias a que nos hallamos de muchas galaxias, con preferencia de las más remotas.

El método habitual para medir la distancia consiste en buscar, en el interior de las galaxias, los objetos cuyo brillo intrínseco esté bien determinado. Los objetos cósmicos más fiables a este propósito son las variables Cefeidas, estrellas cuya tasa de pulsación luminosa depende de su tamaño y, por ende, de su brillo intrínseco. Por desgracia, las Cefeidas son tan débiles que las que pueden distinguirse desde observatorios terrestres pertenecen a lo sumo a una docena de galaxias.

Por ello, los astrónomos se han visto forzados a recurrir a otros objetos: algunas estrellas variables (RR) y determinado tipo de explosiones estelares o supernovas. Se infiere la distancia también de la correlación entre el tamaño (y, por tanto, el brillo) de las galaxias y la velocidad de las estrellas que la forman.

Propósito fundamental del *Telescopio Espacial Hubble* era el de estrechar el margen de las estimaciones observando Cefeidas en galaxias más alejadas. Por culpa del fallo de las lentes primarias se ha postergado esa misión. Aunque podrían suplirla, hasta cierto punto, las mediciones recientes desde el suelo; así, el conjunto de mediciones, sacadas de observaciones de nebulosas planetarias, nubes esféricas de gas emitidas por estrellas vie-

jas, que han realizado Jacoby y Robin Ciardullo, de los Observatorios Nacionales de Astronomía Optica, y Holland C. Ford, de la Universidad Johns Hopkins. Las nebulosas planetarias tienen un espectro característico que puede separarse de la radiación obtenida de las galaxias lejanas. Poseen, asimismo, un máximo para cierto nivel de brillo. Las nebulosas más luminosas sirven de patrón de referencia.

Jacoby y sus colegas midieron las distancias de seis galaxias en el cúmulo de Virgo, congregación de miles de galaxias allende la constelación del mismo nombre. Llegaron a la conclusión de que el cúmulo quedaba bastante cerca, a unos 48 millones de años luz. Ese guarismo cifra la edad del universo en torno a los 12.000 millones de años.

De forma independiente se ha obtenido un acuerdo parcial con esta cifra mediante la técnica desarrollada por John Tonry, del Instituto de Tecnología de Massachusetts; apóyase en que la dificultad implicadas en detectar estrellas individuales e incluso grupos crece con la distancia. El grado de textura detectable de una galaxia indica su lejanía.

Según Tonry, cuando aplica su método a M32, galaxia cercana y de ubicación conocida, se obtienen distancias del cúmulo de Virgo que sólo son un poco mayores que las recabadas por el grupo de Jacoby. Pero añade un grano de duda: al trabajar con un modelo informático de una galaxia, más representativo quizá que la M32, la técnica ofrece distancias a Virgo mucho mayores y congruentes con los 20.000 millones de años de la edad del universo.

La distancia al cúmulo de Virgo resuelta por los grupos de Tonry y Jacoby no discrepa del valor aportado por R. Brent Tully, de la Universidad de Hawai, y Michael J. Pierce, del Observatorio Astrofísico Dominio en la Columbia Británica. Tully y Pierce dedujeron las distancias con la técnica de Tully-Fisher, desarrollada durante los años setenta por Tully y J. Richard Fisher, del Observatorio Nacional de Radioastronomía. El método le saca partido a la tasa de rotación de una galaxia espiral, que aumenta con el tamaño y puede determinarse conociendo las desviaciones de su corrimiento hacia el rojo.

Al analizar las imágenes de cinco galaxias próximas obtenidas con detectores electrónicos finísimos, Wendy L. Freedman, de la Institución Carnegie, llegó a las mismas distancias registradas por Tully y Pierce. Freedman confiesa sentirse todavía insegura de sus resultados, pues necesita calcular la distancia de unas cuantas galaxias más para afirmar con cierto rigor su coincidencia con las distancias calculadas por Tully y Pierce para el cúmulo de Virgo y con la edad del universo. Sin embargo, sus resultados más fiables se muestran coherentes con la antigüedad del universo de 12.000 millones de años.

Allan Sandage, del Observatorio de Monte Wilson, y firme partidario de un universo más antiguo, rechaza las distancias de las galaxias obtenidas por Jacoby, Tully y otros. Para él, "ninguno de estos jóvenes sabe lo que significa la astronomía estadística". Advierte que las observaciones de las supernovas, una de las referencias más fiables para distancias grandes según Sandage, dan estimaciones que concuerdan con la edad del universo de 20.000 millones de años.

John Huchra, de la Universidad de Harvard, mantiene también su propio punto de vista. Después de estudiar muchos años el asunto, cree que las incertidumbres que aparecen al determinar las distancias de las galaxias han sido infravaloradas y que podría ocurrir incluso que la edad del universo oscilara entre los 8000 y los 22.000 millones de años. Sólo el *Telescopio Espacial Hubble*, si volviera a funcionar correctamente, podría estrechar tan amplio margen.

De cualquier modo, muchos astrónomos empiezan a inclinarse por un rebajamiento de la edad del cosmos y ya dan vueltas a sus posibles implicaciones en el campo de la teoría. El mayor problema asociado a un universo más joven reside en su contradicción con la edad asignada a los cúmulos globulares, densos empaquetamientos de estrellas.

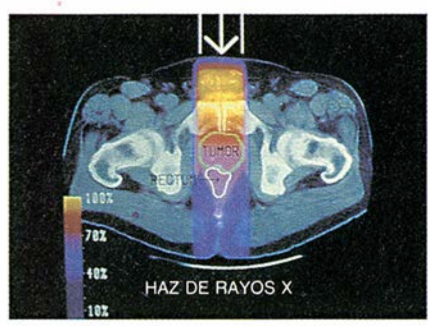
El análisis de la composición química de las estrellas de tales cúmulos abonan una antigüedad mínima de 13.000 millones de años.

El conflicto se endurece cuando se trata de modelos cósmicos que incorporan dos teorías muy relacionadas y prestigiosas: la de un universo inflacionario y la de la materia oscura. De acuerdo con la teoría del universo inflacionario, éste, al poco de su génesis, experimentó un enorme crecimiento. Predice también que el universo posee cantidades ingentes de materia invisible, o materia oscura, que va frenando paulatinamente su expansión.

A tenor de ese cañamazo teórico, las galaxias se alejaban de la Tierra antes más deprisa que hoy, por cuya razón las velocidades de recesión pro-

medio en el tiempo han sido mayores de lo que indican sus corrimientos actuales hacia el rojo. Al dividir estas mayores velocidades de recesión por las distancias menores de Jacoby y otros, se obtiene una edad del universo de entre los 6000 y los 7000 millones de años, en absoluta discrepancia con las edades de los cúmulos globulares.

Antagonismo que puede soslayarse recurriendo de nuevo a la constante cosmológica. Remite ésta a una hipotética energía que llena todo el espacio vacío. Dicha energía residual, cuando se introduce en los modelos cosmológicos, aumenta las estimaciones tanto del tamaño como de la antigüedad del universo. Los teóricos han desdeñado la constante porque complica sus modelos y carece de un soporte empírico, pero empieza ya a considerarse la solución más atractiva de muchos problemas.







3. Haces de protones: depositan su energía directamente en un tumor de próstata. Los rayos X pierden casi toda su energía. En Loma Linda se utiliza una estructura en forma de grúa para dirigir el haz hacia el área a tratar.

Un rayo de esperanza

El hospital clínico de la Universidad Loma Linda, al sur de California, no se amilana ante los proyectos ambiciosos. En 1984 se hizo famosa por el trasplante de un corazón de mandril a un niño de corta edad. Aunque la operación fracasó, el hospital ganó reputación como centro especializado en esa rama de la pediatría quirúrgica.

Loma Linda se prepara ahora para dar otro salto espectacular: posee la máquina terapéutica más costosa del mundo, un acelerador de protones para la terapia del cáncer, bautizado como “el rayo de la esperanza”. Al tratar a un paciente de un melanoma ocular en octubre del año pasado, la máquina de Loma Linda añadía una nueva página al libro de los récords. Incluyendo los costos de los sistemas de visualización y del edificio donde se aloja el acelerador, el monte total del conjunto llega a los 620 millones de pesetas, una tercera parte de los cuales están sufragados por el Departamento de Energía de los Estados Unidos. El hospital ha lanzado una campaña para amortizar las dos terceras restantes.

Los responsables del hospital y los expertos en radiología oncológica afirman que el diseño del acelerador permite tratar más pacientes con haces más precisos y, en algunos casos, más penetrantes que las máquinas existentes de terapia de protones. Estos ingenios se crearon para investigaciones nucleares; los experimentos médicos llegaron más tarde. Desde 1961, el laboratorio del ciclotrón de Harvard ha tratado a 5000 pacientes; desde 1954 un acelerador de partículas experimental del Laboratorio Lawrence de la Universidad de Berkeley ha utilizado iones de helio para conseguir un efecto terapéutico similar en 2000 pacientes. Se quiere, en efecto, llevar al hospital el progreso ya logrado en otras ramas de la investigación científica.

Muchos de los cánceres superficiales que se han tratado con esa terapia de protones han respondido de manera positiva. Los cinco años de tiempo medio de supervivencia para quienes sufren cánceres en la base del cráneo (cordoma y condrosarcoma) lo alcanzan el 70 por ciento de cuantos se someten a esa cura, el doble de los tratados con rayos X.

La técnica del acelerador de Loma Linda es una traducción directa de la empleada en máquinas que trabajan con quarks y muones. El sincrotrón de 50 toneladas, fabricado y diseñado por el Laboratorio Nacional Fermi, despoja, de los electrones, al átomo de hidrógeno antes de inyectar los protones que quedan en el acelerador. Los protones son acelerados por el anillo de 6 metros de diámetro hasta alcanzar energías de entre 70 millones a 250 millones de electronvolts antes de canalizarse en cuatro salas de tratamiento.

Tres de las cuatro salas contienen unas estructuras armadas (grúas) de unos diez metros de altura y 90 toneladas de peso. Alojado en una cavidad en forma de tubo, el paciente debe mantenerse en pie dentro de un molde del cuerpo hecho de plástico y de fibra de vidrio. La grúa gira alrededor del paciente hasta colocarse con el ángulo deseado. Los cuatro imanes dirigen el haz hasta enfocarlo en el blanco durante un minuto o algo más. (El paciente con melanoma recibe un tratamiento con un haz fijo, menos complejo, en la cuarta sala de tratamiento.)

Al variar la energía y la cantidad de protones acelerados, el haz de Loma Linda podrá barrer, línea a línea, una fina capa de tumor; la radiación podrá depositarse capa a capa. Lo mismo que los rayos X, los protones destruyen el tejido fraccionando las moléculas de ADN o formando radicales libres de las moléculas de agua. Pero los rayos X, al entrar en el cuerpo, depositan la radiación ionizante en los primeros centímetros de tejido, y no se detienen al llegar al emplazamiento del tumor. No sucede así con los protones, que, por contra, liberan su energía sólo cuando se frenan hasta pararse, dando lugar a altas dosis de radiación y dejando casi intacto el tejido circundante.

Con el equipo diseñado para el hospital Loma Linda, se cree que será posible tratar los tumores con mayor eficacia que lo hacen los dispositivos experimentales. Los representantes del hospital prevén aplicar esa terapia al cáncer de pulmón, de páncreas y de otros órganos.

Ciencia y empresa

¿Mejor con azúcar?

El corazón no se anda con rodeos en cuestiones circulatorias. Sin sangre que aporte oxígeno, se degradan las moléculas productoras de energía de las células cardiacas. Incluso después de restablecerse la circulación tras un ataque o intervención cardioquirúrgica, la recuperación es lenta. El corazón tarda unos diez días en recobrar sus reservas de energía.

John E. Foker, cirujano cardiovascular del Hospital Clínico de la Universidad de Minnesota en Minneapolis, cree haber hallado la causa del retraso. Cuando se retira la sangre del corazón, las moléculas básicas de la energía celular, las de adenosín trifosfato (ATP), se ven obligadas a romper sus enlaces de alta energía para degenerar en monofosfatos. "Creíamos que se limitaban a quedarse en su sitio, a la espera de ser recargadas", observa Foker. "Pero no es así". Tales metabolitos se deshacen todavía más y salen del corazón, obligándole a sintetizar de nuevo ATP a partir poco menos que de cero.

Foker decidió volver a colocar las piezas en su sitio, "cebar el corazón" para que produjera ATP. Cuando se administró una de ellas, la ribosa, a perros sometidos a intervención quirúrgica, los niveles de ATP recuperaron su concentración normal en sólo 24 horas. Ocurre así porque las células pueden utilizar la ribosa directamente, eludiendo la conversión normal, pero más lenta, de glucosa en ribosa. Sin embargo, hay que administrar ribosa durante cinco días completos. Foker obtuvo una patente norteamericana en 1988 para este "método de estimular la recuperación de la isquemia", anemia local producida por la falta de sangre.

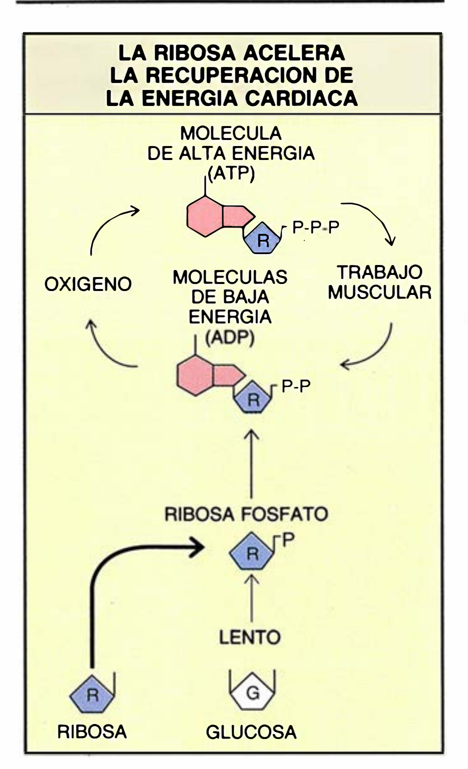
Hay también pruebas esporádicas de la bondad de la ribosa para los humanos. "Muchos de los niños que consiguieron salir del hospital, no habrían podido hacerlo de otra manera", dice Stanley Einzig, jefe de cardiología pediátrica en la Universidad de Virginia en Morgantown. Administró azúcar por vía intravenosa a niños en cuidados intensivos para complementar la terapia convencional de fármacos. "Se puede considerar la ribosa como un suplemento de la nutrición", observa. El azúcar se estudió en el decenio de 1950 para su posible uso en diabetes y no mostró efectos tóxicos.

Pese a ello, no es tan fácil obtener la aprobación de la Oficina norteamericana de Alimentos y Medicamentos (FDA) para tratar a los 300.000 pacientes que se someten a operaciones a corazón abierto o a algunos de los 500.000 que sobrevien a ataques cardíacos cada año. "No basta con poner de manifiesto la recuperación de los niveles de ATP", declara Raymond J. Lipicky, director de la división de productos farmacéuticos cardiorrenales de la FDA. "Tienen que demostrar que el fármaco mejora la mortalidad o la morbilidad: que muere menos gente o se encuentra mejor".

Para demostrar su eficacia se requerirá nada menos que un nuevo método de medir la función cardiaca, calcula Foker. Así que la pequeña compañía que formó para comercializar la ribosa ha pasado los dos últimos años elaborando programas informáticos especializados que corroboren sus observaciones. La firma —Bioenergy, Inc.— está desarrollando una serie de programas para interpretar, con pormenor, los datos de ecocardiogramas. Si el sistema da resultado, es probable que encuentre aplicaciones mucho más allá de la evaluación de la ribosa.

Bioenergy se propone ofrecer registros precisos de la diástole, o capacidad de las cámaras cardiacas para dilatarse y llenarse de sangre. "La diástole es la función cardiaca cuya medida sería más útil", comenta David J. Skorton, profesor de medicina interna de la Universidad de Iowa. Explica que la falta de aportación de nutrientes altera las propiedades de la diástole cardiaca "antes incluso de efectuarse la compresión, o sístole". La presión sanguínea y la circulación de hormonas, factores muy cambiantes, sesgan fácilmente la lectura de la sístole.

Por el momento, los medios finos de determinar la función diastólica son demasiado complicados para resultar prácticos en cualquier sitio que no sea el laboratorio. Los métodos cuya sencillez de uso recomendaría su aplicación en unidades de cuidados intensivos o en clínicas dan poco más que la velocidad a que el corazón se



está llenando. "Esperamos dominar esos extremos" para crear un método de diagnóstico inocuo y científicamente satisfactorio, explica David C. Homans, cardiólogo de la Universidad de Minnesota y encargado de programas de Bioenergy. La compañía está ahora comparando sus programas de ecocardiogrametría con los métodos establecidos de investigación mediante ultrasonidos. Homans quiere medir el espesor y la rigidez de las paredes, la geometría de las cámaras y el volumen cardíaco total, entre otros.

El conocimiento de esas características ayudará a los cardiólogos a realizar comparaciones afinadas del comportamiento de un fármaco —la ribosa, por ejemplo— en un paciente dado y en individuos con diferentes dolencias o medicaciones cardiacas.

El magnetismo de la manta

Los grandes fabricantes de mantas eléctricas de los EE.UU. están tratando de conjurar el peligro que se cierne sobre su negocio. Hasta los periódicos han hablado de los efectos de los campos electromagnéticos sobre la salud. También el gobierno ha tomado cartas en el asunto. Dieciocho miembros del Congreso han pedido que la FDA obligue a hacer público el peligro potencial, para niños y mujeres embarazadas, de los campos eléctricos y magnéticos de las mantas.

Tras la preocupación de los parlamentarios se encuentran estudios que sugieren que los campos producidos por mantas y motores eléctricos pueden provocar cáncer y otras mutaciones celulares. Los usuarios de mantas eléctricas pueden, en el curso de un año, estar sometidos a un campo magnético medio de entre 0,5 y un miligauss, comparable con la exposición media de campos procedentes del conjunto de las demás fuentes, dice H. Keith Florig, quien está estudiando los campos de las mantas subvencionado por el Instituto de Investigación de Energía Eléctrica.

Hasta ahora, los datos sobre los efectos de tales campos, aunque preocupantes, no son definitivos. Un estudio reciente de David A. Savitz y colaboradores de la Universidad de Carolina del Norte en Chapel Hill, publicado en *American Journal of Epidemiology*, indicaba que los hijos de madres que habían usado mantas eléctricas durante el embarazo tenían una probabilidad de padecer tumores cerebrales dos y media veces mayor, así como una probabilidad un 70 por ciento mayor de desarrollar leucemias. El artículo señalaba un aumento total del 30 por ciento en la tasa de cánceres infantiles.

No obstante, el mismo número de la revista publicaba los resultados de las investigaciones de René Verreault y sus colaboradores, del Centro Fred Hutchinson de Investigación del Cáncer en Seattle. No había relación entre el cáncer testicular y el uso de mantas eléctricas en varones adultos. Otros estudios han mostrado resultados igualmente contradictorios.

Aun cuando el jurado esté todavía deliberando, los usuarios han comenzado a reaccionar. Las asociaciones de consumidores aconsejaron que los niños y las mujeres embarazadas prescindiesen de mantas eléctricas y

calentadores de colchones, sustituyéndolos por edredones. Y, después de mantenerse constantes durante muchos años, las ventas de mantas eléctricas bajaron. Los fabricantes atribuyen esa caída a los inviernos templados o a una mala temporada financiera de los minoristas.

Sea el tiempo o el temor lo que afecta a las ventas, algunos fabricantes de mantas están tomando medidas para reducir los campos magnéticos que sus productos generan. En pruebas iniciales realizadas por la FDA, una manta rediseñada, fabricada por Northern Electric Company, registró un campo magnético máximo de entre cinco y diez miligauss en medidas en contacto directo. El fabricante explica que eso equivale a dividir por 20 la intensidad del campo de su producto anterior.

Northern Electric logró anular la mayor parte del campo haciendo que la corriente circulara en sentido opuesto en dos alambres paralelos. Otro fabricante —Casco-Belton Corporation de Grover, Carolina del Norte— también adoptó el mismo principio. Los ensayos provisionales de la FDA mostraron que el rediseño producía también un campo entre 5 y 10 miligauss, entre cinco y diez veces menos del que la compañía declaraba en su antiguo producto.

Estas medidas, sin embargo, no eliminan los campos eléctricos, de los que se ha demostrado en algunos estudios que también causan cambios biológicos. Para esos campos, la FDA dice que puede indicar a los fabricantes que encierren los cables en una pantalla conectada a tierra, lo que representa un cambio de diseño más complicado y costoso que los necesarios para reducir el campo magnético.

Recesión económica

...estas crisis decenales [en los negocios] dependen realmente de variaciones meteorológicas de un período similar, que a su vez dependen, con toda probabilidad, de variaciones cósmicas, de las que son prueba la frecuencia de manchas solares, auroras y perturbaciones magnéticas."

—WILLIAM S. JEVONS, 1879

Los economistas no miran ya el curso de los astros para determinar si la economía de una nación se precipita hacia una recesión. Pero los observadores norteamericanos sí parecen necesitar otear más allá de los mares para calibrar las variaciones económicas. Aunque cada nación es responsable de las grandes oscilaciones de su economía, hacia arriba y hacia abajo, el curso preciso de la variación puede depender bastante de cómo anden los negocios en otras partes.

Durante decenios, los expertos han construido modelos complicados con la esperanza de determinar el mecanismo que rige los ciclos económicos (expansiones y contracciones de una economía). Se han explorado las relaciones entre el producto nacional bruto, el empleo, la liquidez y docenas de otras variables. Mientras predominó el empuje de los EE.UU., otras economías industrializadas desempeñaron sólo un papel marginal en la orientación de los ciclos económicos norteamericanos. Ciertos planes abiertamente orquestados, como los intentos de manejar los tipos de cambio, llevaron a los EE.UU. a un mayor sincronismo con los ciclos económicos de otras naciones. Algunas conmociones económicas globales (principalmente la brusca subida de los precios del crudo en el decenio de 1970) afectaron simultáneamente a muchas economías.

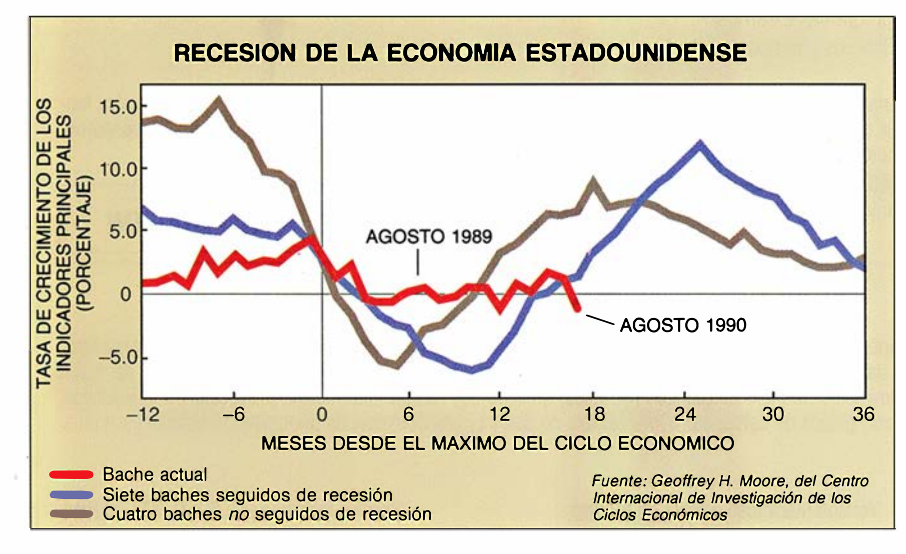
Tradicionalmente, la economía norteamericana ha servido como indicador principal, o señal de aviso anticipado, de las ondas económicas, como observa Geoffrey H. Moore, que dirige el centro de investigación de ciclos económicos internacionales de la facultad de economía de la Universidad de Columbia. Al inicio de una recesión en Norteamérica siguieron meses depresivos más tarde en Europa y Asia. Moore y sus colaboradores recuerdan que el crecimiento de los EE.UU. comenzó a frenarse en febrero de 1989; otros nueve países industriales mostraron signos de debilitación en el pasado verano (Alemania Occidental fue la principal excepción, añade Moore).

A lo largo de los últimos diez años, las economías y los mercados financieros de Japón, Alemania Occidental y otros de los más importantes interlocutores comerciales se han ligado mucho más estrechamente con los norteamericanos, diluyendo así el liderazgo de esta nación. Las importaciones alcanzaron casi el 9 por ciento del producto nacional bruto de los EE.UU. en 1988, por ejemplo, habiendo sido inferiores al 4 por ciento en 1970. (Las exportaciones en 1988 fueron de otro 6,6 por ciento del PNB).

Además, las inversiones extranjeras en los EE.UU., superiores a 72.000 millones el año pasado, importan más de una sexta parte de todas las inversiones nacionales. Antes de 1979, el sobrante del ahorro financiaba las inversiones de capital en el extranjero; el dinero que viene de fuera se considera ahora un elemento principal para el desarrollo de los EE.UU.

Como resultado de ello, las dificultades que se atraviesen fuera del país podrían golpear duramente la ya tambaleante economía norteamericana. Que no haya crecimiento aquí significa que el valor del dólar bajará frente a otras monedas, haciendo así los productos de Cleveland baratos en Cádiz. No obstante, si las otras economías se debilitan también, los fabricantes norteamericanos continuarán sin encontrar muchos clientes.

Si la perspectiva comercial es oscura, la del mercado financiero es negra. Una tumultuosa caída en la bolsa de Tokio, que perdió casi un 40 por ciento de su valor en los diez primeros meses de 1990, ha retraído las inversiones japonesas en los Estados Unidos. En los últimos meses de 1989, las instituciones japonesas adquirieron



17.500 millones de dólares en valores norteamericanos. En cambio, esas instituciones se convirtieron en vendedores netos de tales valores en el primer semestre de 1990, con una cifra de 8900 millones.

Los lazos más estrechos con otras economías significan también que los EE.UU. han sacrificado parte del control sobre su propio destino económico, dice Paul R. Krugman, economista del Instituto de Tecnología de Massachusetts. La Reserva Federal ha dispuesto de enorme poder para estimular una recesión —o detenerla— alterando los tipos de interés. Por ejemplo, la Reserva Federal podía mantener bajo control una recesión reduciendo los tipos de interés, y estimulando así la demanda. Ahora, sin embargo, semejante medida podría debilitar un dólar anémico al encarecer las importaciones y exacerbar la inflación, acelerando posiblemente la huida de capital extranjero.

Una recesión en los EE.UU. dañaría también a otras economías, en particular la de los países en vías de desarrollo. "Si se debilita la capacidad de adquisición de los EE.UU. en bienes duraderos o no —tales como materias primas de Iberoamérica y Asia—, la vitalidad de nuestros clientes se reducirá también", dice Philip Braverman, economista en jefe de DKB Securities de Nueva York, que es una división de Dai Ichi Kangyo Bank.

Otros países podrían sentir menos el efecto. La avidez norteamericana por productos japoneses, tocadiscos y televisores por ejemplo, disminuiría sin duda en una recesión. Pero como el gobierno japonés empezó a ganarse a los consumidores hace tres años para que gastaran en vez de ahorrar, los fabricantes japoneses se han encontrado con hambrientos compradores domésticos, dice Braverman.

Con independencia de los lazos enmarañados que les liguen con otras economías, la responsabilidad final de los espasmos económicos de los EE.UU. es suya y de nadie más. Rudiger Dornbusch, economista del MIT, comenta: "Tenemos unos tipos de interés demasiado altos, un déficit presupuestario excesivo, una industria anticuada, una investigación lentísima, una preparación cada vez menor, una mediocridad sin límites en nuestros gobernantes: todo eso contribuye a una economía lenta. La única industria que ha crecido ha sido la de servicios financieros, y no ha hecho más que meternos en un lío. Pasaremos el próximo decenio pagando esas facturas." Cualquiera juraría que está hablando de España.

El significado de los sueños

Los sueños podrían reflejar un aspecto fundamental del mecanismo de memorización de los mamíferos. ¿Reprocesamos, durante el sueño, información crucial adquirida en estado de vigilia?

Jonathan Winson

A lo largo de la historia, el hombre se ha esforzado por descifrar el significado de los sueños. Los antiguos egipcios les atribuían un poder oracular; la Biblia nos cuenta, por ejemplo, que José vio en los sueños del Faraón el anuncio de siete años de hambre. Otras culturas los han interpretado como fuente de inspiración y de curación o como realidad alternativa.

A lo largo de los últimos cien años, los científicos han ofrecido explicaciones antagónicas de los sueños: las de carácter psicológico y las de tipo neurológico. En 1900, con la publicación de *La interpretación de los sueños,* Sigmund Freud proponía que los sueños constituían la "calzada real" para llegar al subconsciente y revelaban, de manera disfrazada, los más profundos elementos de la vida interior de las personas. Contra esa afirmación aparece la postura más reciente según la cual los sueños carecen de significado; serían el resultado de la actividad aleatoria de las neuronas. El sueño, también se ha dicho, constituiría el mecanismo mediante el cual el cerebro se desharía de la información inútil; vendría a ser un proceso de "aprendizaje a la inversa", de desaprendizaje.

Basándome en mis propios hallazgos y en los obtenidos en otros laboratorios neurológicos, propongo que los sueños tienen un significado real. Los estudios sobre el hipocampo (una estructura cerebral crucial en los procesos relacionados con la memoria), sobre la fase del sueño de movimientos oculares rápidos (REM) y sobre las ondas cerebrales llamadas ritmo theta sugieren que los sueños intervienen de forma destacada en la memorización. En concreto, los estudios del ritmo theta en mamíferos no primates han proporcionado una clave evolutiva para entender el significado de los sueños. Parecen ser el registro nocturno de un proceso mnémico fundamental en los mamíferos: el mecanismo por medio del cual los animales elaboran estrategias para sobrevivir y evalúan sus experiencias habituales a la luz de tales estrategias. La existencia de ese proceso podría explicar el significado de los sueños en el hombre.

JONATHAN WINSON cursó la carrera de ingeniería aeronáutica en el Instituto de Tecnología de California, que acabó en 1946. Se doctoró luego en matemáticas por la Universidad de Columbia y se dedicó a los negocios durante 15 años. Mas, llevado por su interés por las neurociencias, volvió a la universidad, a la Rockefeller, para investigar sobre el procesamiento de la memoria durante la vigilia y el sueño. Desde 1979 da clases allí mientras continúa su trabajo sobre la memoria y el sueño.

En 1953 se desentrañó la fisiología de la ensoñación, con la descripción del ciclo del sueño humano por un grupo de investigadores. En las personas, hallaron, el sueño se inicia con el estado hipnogógico, un período de varios minutos de duración en que los pensamientos consisten en imágenes sueltas o minidramas. Al estado hipnogógico le sigue el sueño de ondas lentas, llamado así porque en ese momento las ondas cerebrales del neocórtex (las circunvoluciones externas del cerebro) son de frecuencia baja y amplitud alta. Esas señales se miden mediante registros electroencefalográficos (EEG).

Se descubrió también que el sueño de la noche procede, asimismo, por períodos discretos en los que las lecturas del EEG son de frecuencia irregular y de amplitud baja, similares a las que se observan en las personas despiertas. Esos períodos de actividad mental configuran el sueño REM. La ensoñación ocurre exclusivamente durante estos intervalos. Las motoneuronas están inhibidas durante el sueño REM, impiden que el cuerpo se mueva libremente y dejan que las extremidades permanezcan ligeramente activas. Los ojos se mueven rápidamente y al unísono bajo los párpados cerrados, la respiración se hace irregular y se acelera el pulso.

La primera fase REM de la noche aparece pasada hora y media de sueño de ondas lentas y dura unos 10 minutos. La segunda y tercera fase REM siguen a episodios más cortos de sueño de ondas lentas y se van dilatando. Tras la fase cuarta y última del sueño REM, que dura de 20 a 30 minutos, el individuo despierta. Si se recuerda el contenido de algún sueño, suele ser el que se tiene en esta última fase.

Al parecer, todos los mamíferos placentarios y marsupiales presentan ese ciclo del sueño, que alterna fases de ondas lentas con sueño REM. En los mamíferos se observan las mismas características asociadas al sueño REM que se presentan en los seres humanos, incluidos los registros de EEG similares a los que ocurren durante la vigilia. Los animales también sueñan. Al destruir las neuronas del tallo cerebral que inhiben el movimiento durante el sueño, los científicos han observado que los gatos dormidos se incorporaban y atacaban o se asustaban por objetos invisibles, imágenes éstas que pertenecían sin duda al mundo de su ensoñación.

En estudios con mamíferos inferiores al grupo de los primates, se han descubierto nuevos aspectos neurofisiológicos del sueño REM. Se ha determinado que el control nervioso de esta fase del ciclo del sueño se encuentra en el tallo cerebral (la región del cerebro más cercana a la médula espinal) y que, durante el sueño REM, las señales nerviosas –las llamadas espigas del córtex ponto-geniculado-occipital (PGO)– proceden

1. ESCALERA DE JACOB, pintada en el año 1973 por Marc Chagall. En ese episodio bíblico, el patriarca sueña que los ángeles suben a los cielos y bajan de ellos a través de una escala.

del tallo cerebral al centro de procesamiento visual, la corteza visual. Las neuronas del tallo cerebral estimulan también la aparición de una onda sinusoidal en el hipocampo. Llámase ritmo theta a esa señal del cerebro.

Hay por lo menos un animal que presenta la fase de ondas lentas y no el sueño REM; no exhibe, pues, el ritmo theta mientras duerme. Se trata del equidna *(Tachyglossus aculeata)*, un mamífero monotrema, ovíparo, que permite ampliar nuestro conocimiento sobre el origen de la ensoñación. La ausencia del sueño REM en el equidna sugiere que esta fase del ciclo del sueño evolucionó hace unos 140 millones de años, cuando los marsupiales placentarios divergieron de la línea de los monotremas. (Los monotremas son los primeros mamíferos que se desarrollaron a partir de los reptiles.)

Considerada desde una perspectiva evolutiva, la perpetuación de un proceso cerebral complejo como el sueño REM indica que nos hallamos ante una función importante para la supervivencia de los mamíferos. El conocimiento de esa función podría revelar el significado en los sueños de los animales dormidos.

Cuando Freud escribió *La interpretación de los sueños*, no se conocía la fisiología del sueño. A la luz del descubrimiento del sueño REM, se modificaron ciertos elementos de su teoría psicoanalítica y se dispuso de una base para elaborar teorías mejor fundadas en la neurología. La ensoñación comenzó a considerarse parte del ciclo del sueño, determinada biológicamente. Ello no obsta para que se siga concediendo crédito a la noción central de la teoría freudiana —es decir, la creencia de que los sueños revelarían nuestros sentimientos y preocupaciones más íntimos—, y se continúe utilizando en la praxis psicoanalítica.

Pero algunos teóricos abandonaron a Freud tras los hallazgos de la neurología. En 1977, J. Allan Hobson y Robert McCarley, de la Facultad de Medicina de Harvard, propusieron la hipótesis de la "activación-síntesis". A tenor de la misma, los sueños consistirían en asociaciones y recuerdos procedentes del prosencéfalo (el neocórtex y estructuras relacionadas) en respuesta a señales aleatorias procedentes del tallo cerebral, como las espigas PGO. Los sueños se limitarían a ser "la mejor respuesta" que el prosencéfalo podría ofrecer a ese bombardeo azaroso que venía del tallo cerebral. Aunque encerraran en ocasiones un contenido psicológico, los sueños se hallarían intrínsecamente carentes de significado.

Hobson ha revisado recientemente su teoría y reconocido el profundo significado psicológico de los sueños. Opina ahora que el significado o plasmación de los sueños resultaría de un orden impuesto sobre el caos de las señales nerviosas: "orden que está en función de nuestra propia visión del mundo, de nuestros recuerdos remotos". En otras palabras, el vocabulario emocional de las personas guardaría relación con los sueños. En una revisión de su hipótesis original más cercana en el tiempo, Hobson ha sugerido también que la activación por parte del tallo cerebral podría servir



de mera conmutación o tránsito de un episodio de ensoñación a otro.

Hobson y McCarley habían aportado una explicación del contenido de la ensoñación, pero seguía sin conocerse la función básica del sueño REM. En 1983, Francis Crick, del Instituto Salk en La Jolla, y Graeme Mitchison, de la Universidad de Cambridge, avanzaron la idea del aprendizaje inverso. Partiendo de la hipótesis de Hobson y McCarley sobre el bombardeo aleatorio del neocórtex mediante ondas PGO, y de su propio conocimiento de la respuesta al estímulo de las redes nerviosas, postularon que la malla de complejas asociaciones nerviosas que constituye el neocórtex podría verse sobrecargada por grandes cantidades de información. El neocórtex podría desarrollar entonces pensamientos falsos, o "parásitos", capaces de comprometer los almacenes verdaderos y ordenados de la memoria.

En el marco de la hipótesis de Crick y Mitchison, el sueño REM serviría para borrar las asociaciones espúreas según una base aplicada con regularidad. Las ondas PGO aleatorias afectarían al neocórtex de forma que se borrara, o se "desaprendiera", la información falsa. Este proceso cumpliría una misión esencial: la de permitir el proceso ordenado de la memorización. En el hombre, los sueños serían un registro operativo de esos pensamientos parásitos o material desechable de la memoria. "Soñamos para olvidar", escriben Crick y Mitchison.

En 1986 ofrecieron una revisión de su hipótesis. La eliminación de los pensamientos parásitos explicaría sólo el contenido extraño de los sueños. Nada podría afirmarse sobre la ensoñación narrativa. Más aún, lo de "soñar para olvidar" se expresaría mejor si escribiéramos "soñar para reducir la fantasía o las obsesiones".

Ninguna de las hipótesis mencionadas explica, a plena satisfacción, la función de los sueños. Por una parte, la teoría de Freud carece de base fisiológica. (Aunque Freud intentó describir la neurología del subconsciente y de los sueños en su "Proyecto para una psicología científica", la empresa resultó prematura y él se ciñó al psicoanálisis.) Por otro lado, a pesar de las adaptaciones de las hipótesis para incorporar elementos psicológicos, las teorías terminaron por negar que los sueños encierren ningún significado.

CORTEX PREFRONTAL
SEPTUM
HIPOCAMPO
CORTEX ENTORRINAL
CORTEX VISUAL
HIPOCAMPO
TALLO CEREBRAL
CIRCUNVOLUCION DENTADA
MEDULA ESPINAL
CELULAS CA1
CELULAS CA3

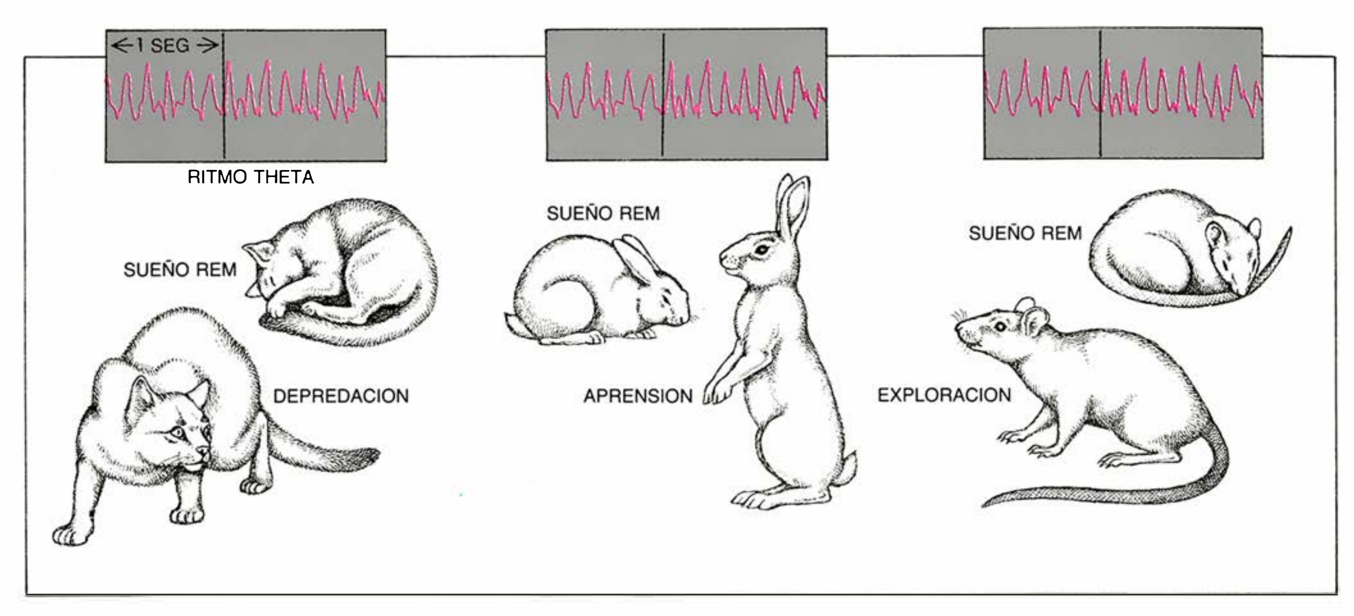
2. ANATOMIA del cerebro y corte transversal del hipocampo. Se nos muestran algunas de las regiones implicadas en la ensoñación. En el hipocampo, se procesa secuencialmente la información procedente del exterior en la circunvolución dentada, las células piramidales CA3 y las CA1. En las especies de mamíferos no primates, el ritmo theta se genera en la circunvolución dentada y las CA1.

En la exploración de los aspectos neurocientíficos del sueño REM y de la memoria me pareció atisbar el mejor camino para adentrarme en el significado y la función de los sueños. La clave de esta investigación estuvo en el ritmo theta.

John D. Green y Arnaldo A. Arduini, de la Universidad de California en Los Angeles, descubrieron, en 1954 y en animales en estado de vigilia, el ritmo theta. Observaron una señal sinusoidal regular de seis ciclos por segundo en el hipocampo de los conejos, cuando éstos se sentían inquietos por estímulos procedentes del exterior. Denominaron a esta señal ritmo theta, en razón de un componente del EEG de igual frecuencia descubierto con anterioridad.

Ese ritmo se confirmaría más tarde en el tupaya, el topo, la rata y el gato. Aunque todos lo mostraban en estado de vigilia, el ritmo theta estaba relacionado con comportamientos muy diferentes de acuerdo con la especie. Así, en señalado contraste con el conejo, los estímulos ambientales no provocaban el ritmo theta en la rata, que sólo lo mostraba durante el movimiento y, de forma característica, cuando se hallaba en exploración. Sin embargo, en 1969, Case H. Vanderwolf, de la Universidad de Ontario, descubrió que había un momento en que los animales que estaba estudiando, entre ellos la rata, presentaban el ritmo theta: el sueño REM.

En 1972 publiqué un comentario en el que señalaba que la aparición dispar del rimo theta podría hallar explicación en el comportamiento del animal. Los individuos despiertos parecían mostrar ritmos theta cuando asumían conductas cruciales para su supervivencia. Es decir, el ritmo theta aparecía cuando desarrollaban un comportamiento que no se hallaba genéticamente determinado —como la alimentación o la conducta sexual—, sino que constituía una respuesta ante la información procedente de un entorno cambiante. El comportamiento depredador del gato, el de presa del conejo y el explorador de la rata son, respectivamente, fundamentales para su supervivencia. Una rata hambrienta explorará antes de comer, aun cuando tenga la comida delante.



3. RITMO THETA presente durante diferentes conductas durante la vigilia en distintas especies. Cada una de esas conductas es fundamental para la supervivencia del animal. En placentarios y marsupiales, el ritmo theta está presente durante el sueño de movimientos oculares rápidos (REM).

Además, como el hipocampo interviene en los procesos mnémicos, la presencia del ritmo theta durante el sueño REM en esa región del cerebro podría estar relacionada con dicha actividad. Propuse que el ritmo theta reflejaba una función nerviosa en virtud de la cual se reprocesaba, en la memoria y durante el sueño REM, información esencial para la supervivencia de la especie, acumulada durante el día.

En 1974, mientras registraba señales del hipocampo en ratas y conejos que podían moverse libremente, descubrí la fuente generadora del ritmo theta en el hipocampo. Se admite que el hipocampo constituye, con el neocórtex, la estructura nerviosa básica para el almacenamiento de la memoria. El hipocampo (así llamado por su forma de "caballito de mar") es una estructura secuencial que se compone de tres tipos de neuronas. La información procedente de todas las áreas sensitivas y de asociación del neocórtex converge en el córtex entorrinal. Desde allí, se transmite hacia las tres poblaciones sucesivas de neuronas del hipocampo. La señal llega en primer lugar a las células granulosas de la circunvolución dentada, después a las células piramidales CA3 (así llamadas por su forma triangular) y, por último, a las células piramidales CA1. Tras el procesamiento de la información por estos tres tipos de células, se retransmite al córtex entorrinal y, de allí, al neocórtex.

Los estudios realizados por mí demostraron que el ritmo theta se producía en dos regiones del hipocampo: la circunvolución dentada y las neuronas CA1. Los ritmos de esas dos áreas eran sincrónicos. Con posterioridad, Susan Mitchell y James B. Ranck, del Hospital Clínico de la Universidad de Nueva York, identificaron un tercer generador sincrónico en el córtex entorrinal; Robert Verdes, de la Universidad estatal de Wayne, describió las neuronas del tallo cerebral que controlan el ritmo theta. Estas neuronas transmiten señales al septum (estructura del prosencéfalo) que, a su vez, activa el ritmo theta en el hipocampo y en el córtex entorrinal. Tenemos, pues, que el tallo cerebral activa el hipocampo y el neocórtex, el sistema mnémico nuclear del cerebro.

Con objeto de determinar la relación entre el ritmo theta y la memoria, lesioné el septum de la rata. Las ratas que con anterioridad habían aprendido, mediante señales espaciales, a localizar una posición particular del laberinto, no eran ya capaces de hacerlo. Sin el ritmo theta, se perdía la memoria espacial.

Los estudios sobre los cambios celulares relacionados con la memoria ilustran el papel del ritmo theta. El descubrimiento de la potenciación a largo plazo (PLP) en 1973 puso de manifiesto, en particular, los mecanismos mediante los que podría codificarse la memoria. Timothy V. P. Bliss y A. R. Gardner-Medwin, del Instituto Nacional de Investigaciones Médicas en Londres, y Terje Lømo, de la Universidad de Oslo, encontraron alteraciones en las células nerviosas a las que se habían aplicado estímulos intensos con pulsos eléctricos.

En investigaciones precedentes se había observado que, si se estimulaba la vía entre el córtex entorrinal y las células granulosas del hipocampo, la respuesta de estas células se podía medir con un electrodo de registro. Sirviéndose de esa técnica, Bliss y sus colaboradores midieron la respuesta normal a un solo pulso eléctrico. Aplicaron luego a la misma vía una larga serie de señales de alta frecuencia –pulsos tetánicos–. Tras la serie de estímulos tetánicos, el pulso eléctrico solitario provocaba una excitación en las células granulosas mucho mayor que la observada antes del experimento. El efecto de refuerzo persistía durante tres días. Este fenómeno de PLP constituía el tipo de aumento de capacidad nerviosa que podía mantener la memoria. La PLP se considera ahora como un modelo del aprendizaje y la memoria.

La aparición de la PLP obedece a la actividad del receptor del NMDA (*N*-metil-D-aspartato). Esta molécula la portan las dendritas de las células granulosas y de las células CA1 del hipocampo, así como las neuronas del neocórtex. Lo mismo que ocurre con otros receptores neuronales, un neurotransmisor —el glutamato en este caso— activa al receptor del NMDA. El glutamato abre momentáneamente un canal diferente del NMDA en la dendrita de la célula granulosa, dejando que el sodio del espacio extracelular penetre en el interior de la neurona. Este flujo induce la despolarización de la célula granulosa. Si dicha despolarización es suficiente, la célula se excita, transmitiendo información a otras neuronas.

A diferencia de otros receptores neuronales, el del NMDA posee una propiedad adicional. Si sucede una

nueva activación por parte del glutamato mientras la célula granulosa está despolarizada, se abre un segundo canal, que permite la entrada de calcio. Se cree que el calcio actúa de segundo mensajero y desencadena una cascada de episodios intracelulares que culminan con cambios de larga duración en las sinapsis o PLP. (Forzosamente hemos simplificado aquí la descripción del proceso; la verdad es que la PLP es objeto hoy de intensa investigación.)

El impulso tetánico aplicado por Bliss y sus colaboradores no ocurría de forma natural en el cerebro. Por eso, la cuestión planteada seguía siendo la consecución de una PLP en circunstancias normales. En 1986, John Larson y Gary S. Lynch, de la Universidad de California en Irvine, y Gregory Rose y Thomas V. Dunwiddie, de la Universidad de Colorado en Denver, sugirieron la relación entre la aparición de la PLP en el hipocampo y el ritmo theta. Aplicaron un pequeño número de pulsos eléctricos a las células CA1 del hipocampo de la rata y produjeron una PLP, pero sólo cuando transcurría entre pulsos el intervalo de tiempo normal que media entre dos ondas theta (unos 200 milisegundos). El ritmo theta constituye, por lo que se ve, el medio natural a través del cual se activa el receptor del NMDA en las neuronas del hipocampo.

El trabajo desarrollado en mi propio laboratorio de la Universidad Rockefeller corroboró los hallazgos de Larson y Lynch obtenidos con las células CA1, si bien en mi caso se trataba de las células granulosas del hipocampo. Con el auxilio de Constantine Pavlides y Yoram J. Greenstein demostré que la PLP dependía de la presencia y la fase del ritmo theta. Si se aplicaban los pulsos eléctricos a las células en el pico de las ondas theta, se inducía PLP. Pero si el mismo pulso se aplicaba en el seno de las ondas —o en ausencia del ritmo theta—, no se producía PLP.

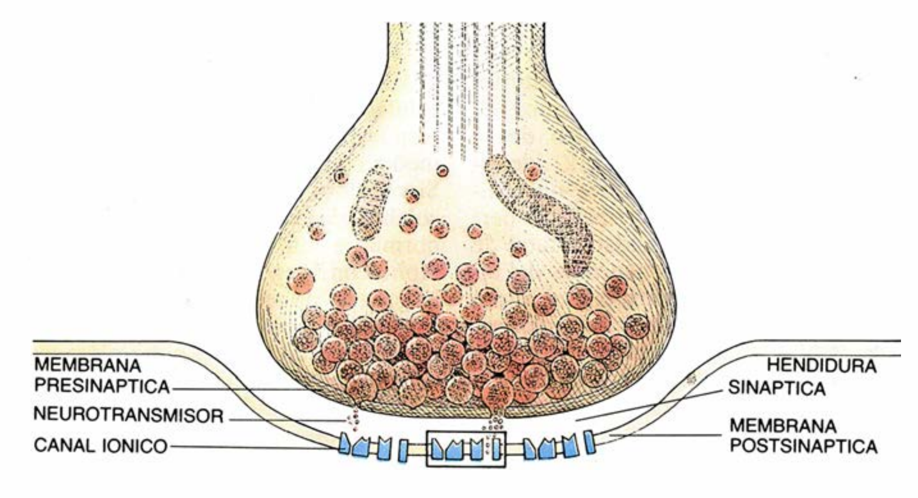
Se iba perfilando un cuadro de explicación coherente del proceso de la memoria. Cuando una rata se halla explorando, las neuronas del tallo cerebral producen un ritmo theta. La información olfatoria (que en la rata está sincronizada con el ritmo theta, como lo está el movimiento de los bigotes) y la de otros estímulos sensoriales convergen en el córtex entorrinal y en el hipocampo. Allí, el ritmo theta fracciona la información en "bits" de 200 milisegundos. Los receptores del NMDA, actuando junto con el ritmo theta, permiten el almacenamiento a largo plazo de esta información.

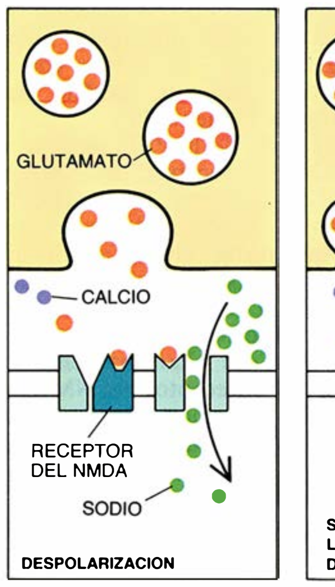
Durante el sueño REM ocurre un proceso similar. Aunque no existe entrada de información o movimiento durante esa fase del sueño, la red formada por el neocórtex y el hipocampo es recorrida por el ritmo theta, capaz éste de causar quizá modificaciones de larga duración en la memoria.

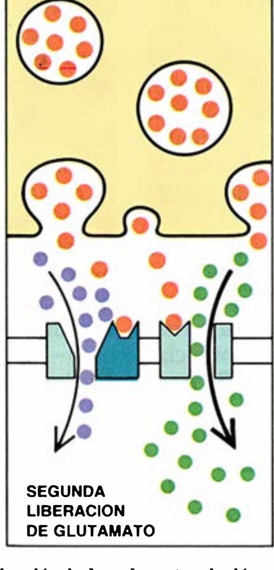
Los resultados de uno de mis experimentos posteriores sirvieron para demostrar que, durante el sueño, se almacenaba la memoria espacial en el hipocampo de la rata. John O'Keefe y J. Dostrovsky, de la Universidad de Londres, demostraron que en el hipocampo de la rata se excitan determinadas neuronas CA1 cuando el animal despierto se mueve hacia cierto espacio: el campo de lugar de las neuronas. Inferíase de este hallazgo que la neurona CA1 se estimulaba para cartografiar el medio exterior y confinarlo a la memoria.

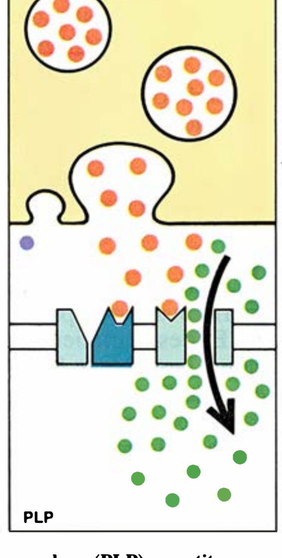
En 1989, localicé con Pavlides dos neuronas CA1 en el hipocampo de la rata que tenían campos de lugar diferentes. Registramos las señales procedentes de ambas células simultáneamente. Tras determinar la excitación normal de las ratas en animales despiertos y dormidos, situamos a una rata en el campo de lugar de una de las neuronas. La neurona se excitó vigorosamente, trazando el mapa de la posición. La segunda célula se excitó, aunque mucho menos, pues no estaba codificando el espacio. Seguimos registrando las señales de ambas neuronas conforme la rata se movía; después entró en varios ciclos de sueño. Así se estudiaron seis pares de neuronas.

Descubrimos que las neuronas que

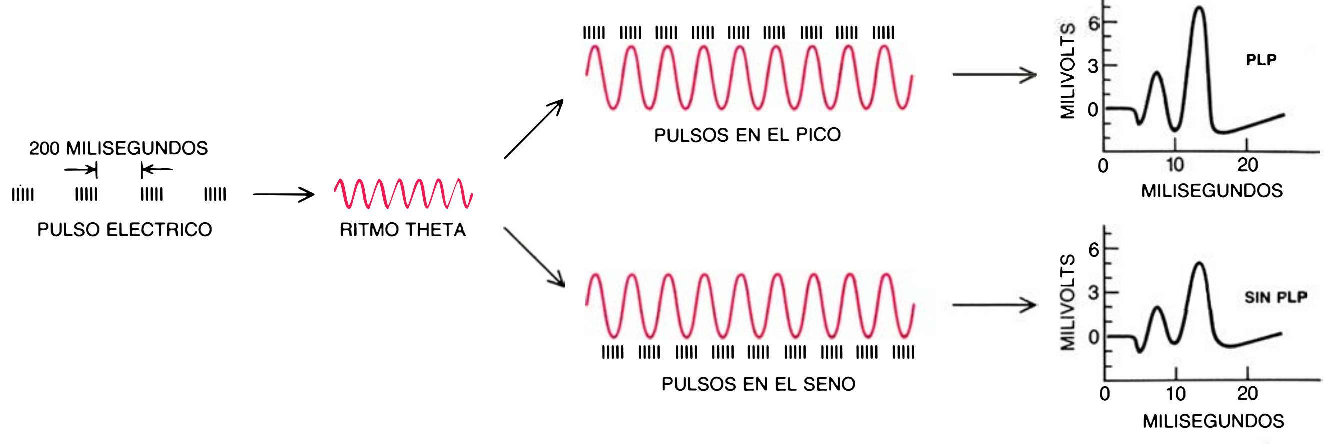








4. RECEPTOR NMDA, cuya activación induce la potenciación a largo plazo (PLP); constituye un modelo de la memoria. La liberación del neurotransmisor glutamato (*cuadro izquierdo*) abre un canal diferente del receptor del NMDA, permitiendo la entrada de sodio, que despolariza la neurona. Si se produce una nueva liberación de glutamato mientras la célula aun está despolarizada (*cuadro central*), el receptor del NMDA abre un segundo canal que permite la entrada del calcio, promoviendo la PLP. Esta potenciación a largo plazo se produce en virtud del aumento del paso del sodio a través del canal independiente del NMDA (*cuadro derecho*) y por la mayor despolarización subsiguiente.



5. PLP en las células granulosas del hipocampo, de la que es responsable el ritmo theta. Los pulsos eléctricos, separados entre sí por 200 milisegundos (el tiempo entre los picos consecutivos de dos ondas theta), aplicados en el pico del ritmo theta, producen una potenciación a largo plazo.

habían codificado el espacio se estimulaban de forma habitual cuando el animal se movía antes de dormir. Sin embargo, durante el sueño, se estimulaban a un ritmo significativamente mayor que el correspondiente a su ritmo basal en sueños anteriores. No se producía tal aumento del ritmo de estimulación durante el sueño en las neuronas que no habían codificado el espacio. Este experimento sugería que, durante el sueño, tenía lugar, en determinadas neuronas, un reprocesamiento o intensificación de la información que el animal había codificado durante el estado de vigilia.

La prueba de que el ritmo theta codifica los recuerdos durante el sueño REM podría ofrecérnosla, no sólo la investigación neurológica, sino también los estudios evolutivos. La aparición de un mecanismo nervioso para procesar la memoria en el sueño REM traza una divisoria en la anatomía cerebral de los animales que presentan ese aspecto del ciclo del sueño y los que no lo presentan. Existen, en efecto, diferencias a ese respecto entre el equidna y los marsupiales y placentarios.

El equidna tiene un gran córtex prefrontal circunvolucionado, con una relación con el resto del cerebro mayor que la que se da en cualquier otro mamífero, hombre incluido. Necesita su enorme córtex prefrontal para realizar, así lo creo, una doble función: reaccionar debidamente ante la información exterior, basándose en la experiencia acumulada, y evaluar y almacenar la información nueva para su supervivencia en el futuro. Sin ritmo theta durante el sueño REM, el equidna no es capaz de procesar la información mientras duerme; sí lo presenta cuando busca comida. Para desarrollar aptitudes más avanzadas, el córtex prefrontal debería haber aumentado progresivamente de tamaño, más allá de la capacidad del cráneo, a menos que se hubiera ingeniado otro mecanismo cerebral.

El sueño REM podría ser ese nuevo mecanismo, permitiendo que el procesamiento de la memoria ocurriera "fuera de contacto". Coincidiendo con el desarrollo ostensible del sueño REM en marsupiales y placentarios, se produjo un cambio anatómico notable: el córtex prefrontal redujo bastante su tamaño. No era tan imprescindible para procesar la información. Por consiguiente, ese área cerebral podía desarrollarse para ofrecer una capacidad cognitiva y perceptiva avanzada en especies superiores.

La naturaleza del sueño REM respalda esa argumentación evolutiva. A lo largo del día, los animales recogen información que conlleva locomoción y movimiento de los ojos. El reprocesamiento de esta información durante el sueño REM no podría separarse fácilmente del movimiento relacionado con la experiencia; tal disociación cabría sólo en el caso de una modificación drástica de los circuitos cerebrales. Por consiguiente, para mantener el sueño, había que suprimir el movimiento mediante la inhibición de las neuronas motoras. La supresión del movimiento ocular es innecesaria, ya que esta actividad no altera el sueño.

El movimiento ocular rápido en el estado de vigilia y durante el sueño REM se ve acompañado de potenciales del movimiento ocular, similares a las espigas PGO. La función de estas señales, todavía por establecer, podría servir para avisar a la corteza visual acerca de la información externa cuando el animal está despierto y reflejar el reprocesamiento de dicha información durante el sueño REM. En cualquier caso, las espigas PGO no alteran el sueño y no es necesario suprimirlas, a diferencia de las neuronas motoras.

Con la evolución del sueño REM, cada especie podía procesar la información de interés para su supervivencia: localización de la comida, depredación o huida, actividades durante las cuales está presente el ritmo theta. En el sueño REM podría volverse a acceder a esa información e integrarla con la experiencia anterior para proporcionar estrategias de conducta.

Aunque no se ha comprobado la existencia del ritmo theta en primates, las señales del cerebro aportan una clave para descifrar el origen de los sueños en humanos. Los sueños podrían reflejar un mecanismo de procesamiento de la memoria heredado de especies inferiores que permitiría, durante el sueño REM, volver a procesar la información útil para la supervivencia. Esta información podría constituir el sustrato del subconsciente.

Como los animales no poseen lenguaje, la información que procesan durante el sueño REM es, por necesidad, sensorial. En consonancia con nuestros primitivos orígenes mamíferos, los sueños del hombre son de naturaleza sensorial, visual en primer lugar. Los sueños no toman la forma de narración verbal.

De acuerdo también con el papel que el sueño REM desempeña en el procesamiento de los recuerdos en los animales, no hay necesidad funcional de que este material pase a la conciencia. Aparece ésta en una fase tardía de la evolución, con el hombre. Pero tampoco existen motivos para que el material de los sueños no llegue a hacerse consciente. Se pueden, pues, recordar los sueños con más facilidad cuando uno se despierta durante el sueño REM, o inmediatamente después.

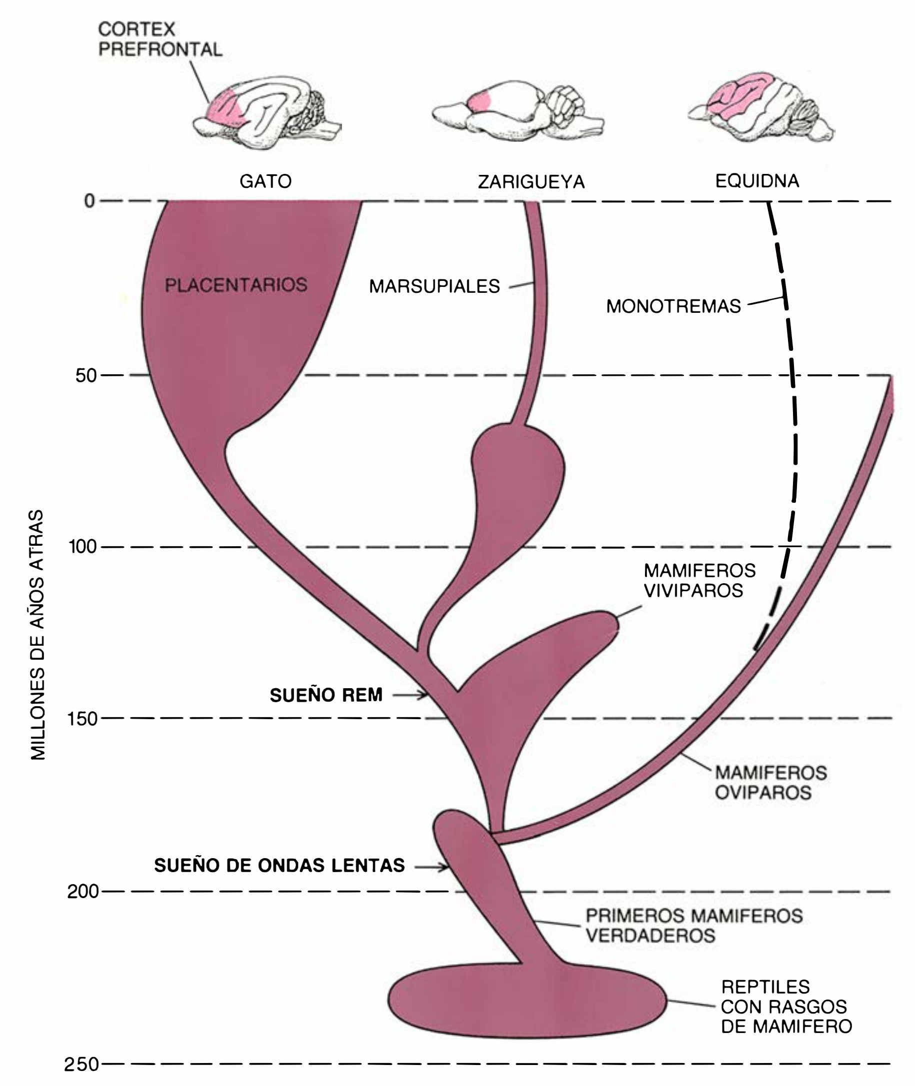
En coherencia con la evolución y con los datos extraídos de la neuro-

ciencia y de los contenidos de los sueños, sugiero que éstos reflejan la estrategia del individuo para sobrevivir. Los temas de los sueños, muy variados y complejos, se tejen con imágenes de uno mismo, sus miedos e inseguridad, fortaleza, ideas grandiosas, orientaciones sexuales, deseos, celos y amores.

Los sueños poseen, sin la menor duda, un sustrato psicológico profundo. Así han opinado los psicoanalistas posteriores a Freud y ha ilustrado Rosalind Cartwright, del Hospital presbiteriano de San Lucas de Chicago. Cartwright estudia un universo de 90 individuos en proceso de separación matrimonial y divorcio. Todas las personas son evaluadas clínica y psicológicamente para determinar sus actitudes y respuestas a su crisis personal. Se les despierta en el período de sueño REM para que cuenten sus sueños, que interpretan ellos mismos, sin ninguna pregunta que pudiera condicionar su explicación. En los 70 individuos estudiados hasta el momento, los sueños expresan sus pensamientos inconscientes; su contenido guarda estrecha relación con las conductas que adoptan para afrontar su crisis durante la vigilia.

A pesar de que el tema "elegido" para considerar durante el sueño de una noche es impredictible, ciertas dificultades de la vida –como en el caso de los sujetos de Cartwright– relacionadas con la supervivencia psicológica son los seleccionados para su procesamiento en el sueño REM. En los períodos ordinarios de la vida, dependiendo de la personalidad del individuo, el contenido de lo soñado puede ser totalmente libre. Además, cuando se une a las complejas asociaciones que forman parte intrínseca del procesamiento durante el sueño REM, el significado de lo soñado puede resultar obscuro.

Sin embargo, existen razones para creer que el proceso cognitivo que ocurre en los sujetos de Cartwright se da en todas las personas. La interpretación del mensaje coherente que se está emitiendo depende de la perspicacia de la persona en relación con hechos similares o importantes. Estas asociaciones acusan el influjo poderoso de las primeras experiencias infantiles.



6. ARBOL EVOLUTIVO que muestra la divergencia entre marsupiales y placentarios, por un lado, y monotremas por otro. El equidna, que no presenta sueño REM, tiene un córtex prefrontal mayor en comparación con el resto del cerebro que cualquier mamífero restante, incluido el ser humano. Es bastante mayor que el de otros animales de tamaño similar, como la zarigüeya y el gato.

Mi hipótesis ofrece también una explicación para la gran cantidad de sueño REM que experimentan los bebés y los niños. Los recién nacidos pasan ocho horas al día en sueño REM. En ese momento de su vida, el ciclo del sueño está desorganizado. El sueño tiene lugar en períodos de 50 a 60 minutos y empieza con el REM, en vez de hacerlo con el sueño de ondas lentas. A los dos años, el sueño REM se reduce a tres horas al día y se establece el patrón del adulto. En adelante, el tiempo dedicado al sueño REM cae gradualmente hasta algo menos de dos horas.

El sueño REM podría desempeñar una función especial en los recién nacidos. Se opina que estimula el crecimiento nervioso. Pero, cualquiera que sea su función, cuando el hipocampo, que se encuentra todavía en fase de desarrollo en el recién nacido, se hace funcional, el sueño REM cumple su misión de memoria interpretativa. La información durante la vigilia que debe integrarse en ese momento del desrrollo constituye el sustrato cognitivo básico de la memoria, el concepto del mundo real a partir del cual se compararán e interpretarán las experiencias posteriores. La organización en la memoria de esta amplia infraestructura requiere ese tiempo adicional de sueño REM.

Por razones que no pudo conocer, Freud expuso en su trabajo una gran verdad. Hay un subconsciente; y los sueños constituyen el "camino real" que lleva a su desciframiento. Ahora bien, las características del subconsciente y los procesos relacionados del funcionamiento del cerebro son muy diferentes de lo que Freud pensaba. En mi opinión, no se trata de una caldera de pasiones desenfrenadas y deseos de destrucción, si no de una estructura mental cohesionada y en continua actividad que toma nota de las experiencias de la vida y reacciona de acuerdo con su propio esquema de interpretación. Los sueños no están camuflados por culpa de la autorrepresión. Deben su carácter insólito a las asociaciones complejas procedentes de la memoria.

La investigación acerca del sueño REM abona la existencia de una razón biológica importante para soñar.

La versión revisada de la hipótesis de la activación-síntesis de Hobson y McCarley reconoce el sustrato psicológico profundo de los sueños. En su actual forma truncada, la hipótesis de la activación aleatoria por parte del tallo cerebral tiene escaso poder explicativo o predictivo.

La hipótesis de Crick y Mitchison otorga una función al sueño REM: el aprendizaje inverso. Ahora bien, eso no puede aplicarse a los elementos narrativos de los sueños, sino tan sólo a los elementos extravagantes. Deben definirse las posibles implicaciones de todo ello en relación con el procesamiento REM en especies inferiores antes de someter a valoración la teoría. Además, la hipótesis de Crick y Mitchison aplicada al hipocampo sugiere que las neuronas se excitan aleatoriamente durante el sueño REM, causando el aprendizaje inverso. Pero en mi experimento con las neuronas que codifican el espacio, dichas neuronas se excitan de forma selectiva, prueba de que se da un proceso ordenado de la memoria.

La investigación habrá de seguir contribuyendo a desentrañar el significado de los sueños. En concreto, sería necesario determinar si la supresión del ritmo theta durante el sueño REM recorta la memoria. La ausencia de ritmo theta en los primates se explica quizá mediante el desplazamiento, en el predominio, del olfato por el sentido de la vista. En el hipocampo podría haber un mecanismo nervioso equivalente que periódicamente activara al receptor del NMDA. Ese tipo de estudios y otros parecidos permitirán conocer aspectos fundamentales del proceso de memorización y del soporte neurológico de la estructura psicológica humana.

BIBLIOGRAFIA COMPLEMENTARIA

LOSS OF HIPPOCAMPAL THETA RHYTHM RESULTS IN SPATIAL MEMORY DEFICIT IN THE RAT. Jonathan Winson en *Science*, vol. 201, n.° 435, págs. 160-163; 1978.

BRAIN AND PSYCHE: THE BIOLOGY OF THE UNCONSCIOUS. Jonathan Winson. Anchor Press, Doubleday, 1985.

LONG-TERM POTENTIATION IN THE DENTATE GYRUS IS INDUCED PREFERENTIALLY ON THE POSITIVE PHASE OF ORHYTHM. Constantine Pavlides, Yoram J. Greenstein, Mark Grudman y Jonathan Winson en *Brain Research*, vol. 439, págs. 383-387; 1988.

INFLUENCES OF HIPPOCAMPAL PLACE CELL FIRING IN THE AWAKE STATE ON THE ACTIVITY OF THESE CELLS DURING SUBSEQUENT SLEEP EPISODES. Constantine Pavlides y Jonathan Winson en *Journal of Neuroscience*, vol. 9, n.° 8, págs. 2907-2918; agosto de 1989.

Teoría de nudos y mecánica estadística

Teorías matemáticas concebidas para la física cuántica forjan un eslabón que une estos dos campos del conocimiento tan dispares entre sí

Vaughan F. R. Jones

En 1984 me tropecé con un conjunto de técnicas que encadenaban entre sí dos de las especialidades más dispares, en apariencia, de las matemáticas y la física, a saber, la teoría de nudos y la mecánica estadística. La mecánica estadística estudia sistemas integrados por números inmensos de elementos componentes; los sistemas pequeños carecen casi siempre de importancia. En la teoría de nudos, por contra, hasta los nudos y enlaces más pequeños poseen sutiles propiedades.

Y no obstante, ciertas relaciones algebraicas utilizadas para resolver modelos de mecánica estadística desempeñaron un papel clave en la descripción de una propiedad matemática de los nudos conocida por un invariante polinómico. Esta conexión, tenue al principio, se ha desarrollado hasta constituir un caudaloso flujo de ideas. El afloramiento de este territorio común no es atípico en los desarrollos matemáticos y físicos recientes; antes bien, de la interacción de ideas procedentes de especialidades diferenciadas resultan frutos importantes e inesperados.

A decir verdad, el descubrimiento de la conexión entre los nudos y la mecánica estadística tuvo por intermediaria una teoría íntimamente relacionada con la estructura matemática de la física cuántica. Dicha teoría, la de álgebras de von Neumann, se distingue por ser en ella la dimensión una variable continua. Lo típico y habitual es que la dimensión de los espacios venga expresada por un número natural, como 2, 3 u 11, pero en las álgebras de von Neumann son igualmente posibles dimensiones como $\sqrt{2}$ o π. La posibilidad de que la dimensión fuese una variable continua desempeñó un papel clave en la unión entre la teoría de nudos y la mecánica estadística.

Al estudiar la situación en sentido contrario, se descubrió sin tardanza que los invariantes de los nudos aparecían también en teoría cuántica de campos. En efecto, Edward Witten, del Instituto de Estudios Avanzados de Princeton (New Jersey), ha demostrado que la teoría cuántica de campos "topológica" constituye una vía natural para la expresión de las nuevas ideas sobre los nudos. Este progreso ha permitido, a su vez, una preciosa generalización de los invariantes de los nudos a espacios tridimensionales más complicados, conocidos por variedades tridimensionales, en los que el propio espacio puede contener huecos y bucles.

La nueva teoría de nudos ha mostrado ya su utilidad en otro dominio completamente desconectado de los anteriores. Se ha establecido en biología molecular que las dobles hélices de ADN se anudan y encadenan en el curso de procesos biológicos como los de recombinación y replicación. Los mecanismos de que se valen las células para desanudarlos guardan una misteriosa semejanza con el más sencillo de los métodos de obtención de los nuevos invariantes polinómicos.

Los nudos han servido para fines prácticos o decorativos desde tiempos inmemoriales. Los marinos han ideado nudos muy elaborados —de nombres no menos prolijos en ocasiones— para sus necesidades. Sólo a fines del siglo XIX comenzaron los matemáticos a interesarse por los nudos. Lord Kelvin, por ejemplo, se esforzó en deducir la estructura de la tabla periódica de los elementos tomando como hipótesis que los átomos eran, en realidad, vórtices anudados en el seno del "éter". (Aunque dicho trabajo resultó infructuoso, sirvió de inspiración a Peter G. Tait para crear las primeras tablas de nudos, donde éstos se enumeraban atendiendo a cierto orden de complejidad.)

La teoría de nudos se ha convertido desde entonces en rama fecunda de las matemáticas. Una de las bellezas de esta disciplina estriba justamente en lo muy familiares que son sus principales objetos de estudio: basta tomar un cordel y unir sus extremos; el bucle así formado constituye un modelo perfectamente adecuado del concepto de "curva regular y cerrada que no se corta a sí misma" utilizado por el matemático. Una versión más general de nudo, llamada enlace, puede constar de más de un bucle. Se dice que dos nudos o enlaces son iguales si resulta posible lograr que ofrezcan exactamente el mismo aspecto tensando o aflojando la cuerda, sin cortarla.

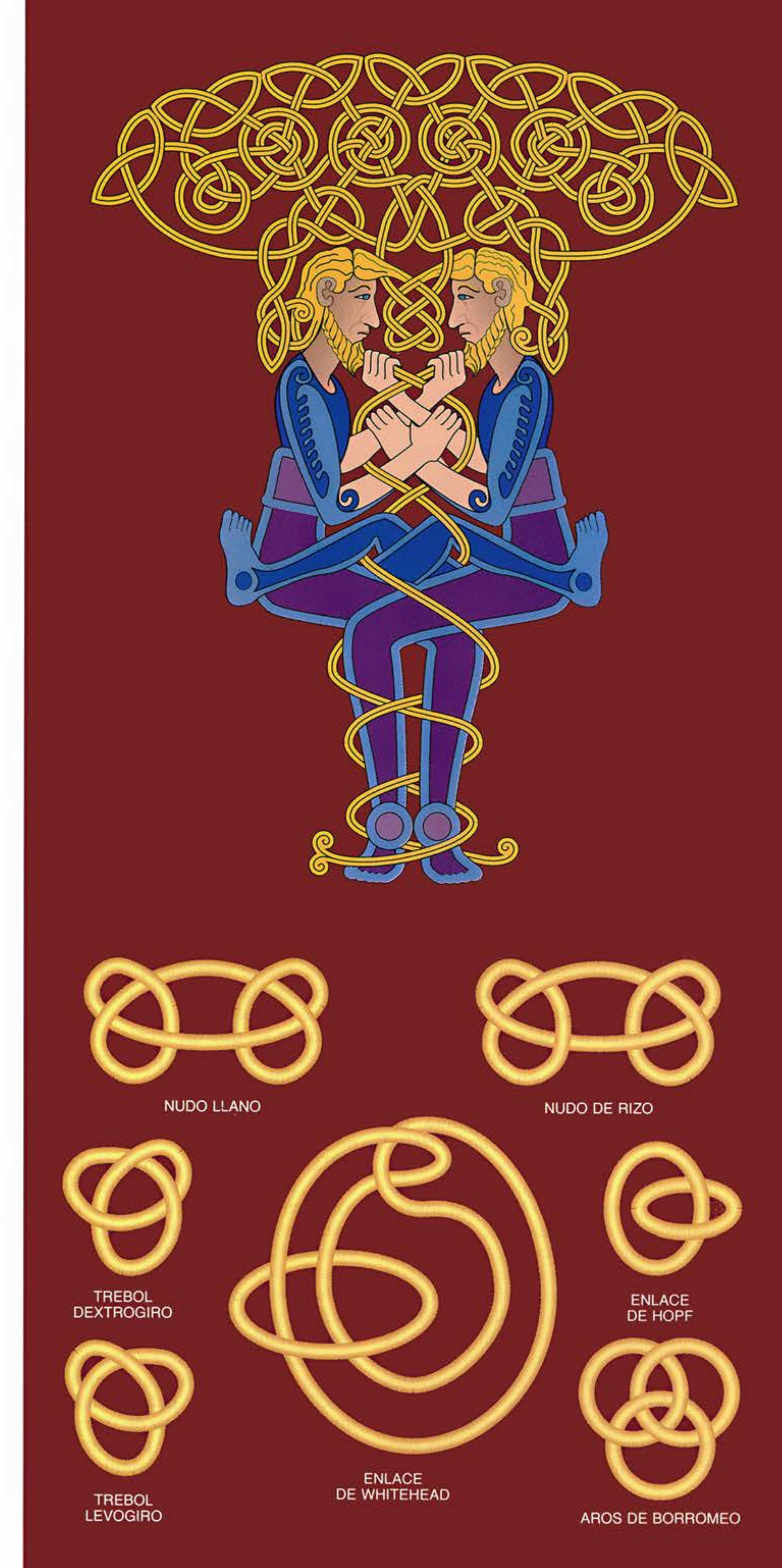
Fijémonos en un sencillo bucle cerrado de cordel, yacente sobre una superficie plana. Inmediatamente saltan a la vista dos importantes características de la teoría de nudos. En

VAUGHAN F. R. JONES explica matemáticas en la Universidad de California en Berkeley. Tras licenciarse por la Universidad de Auckland en 1973, fue a Ginebra, donde prosiguió sus estudios hasta recibir el doctorado en 1979. Simultanea su docencia en Berkeley con estancias en el Instituto de Altos Estudios Científicos de París. En agosto del año pasado le fue concedida la Medalla Fields, equivalente en matemática al premio Nobel.

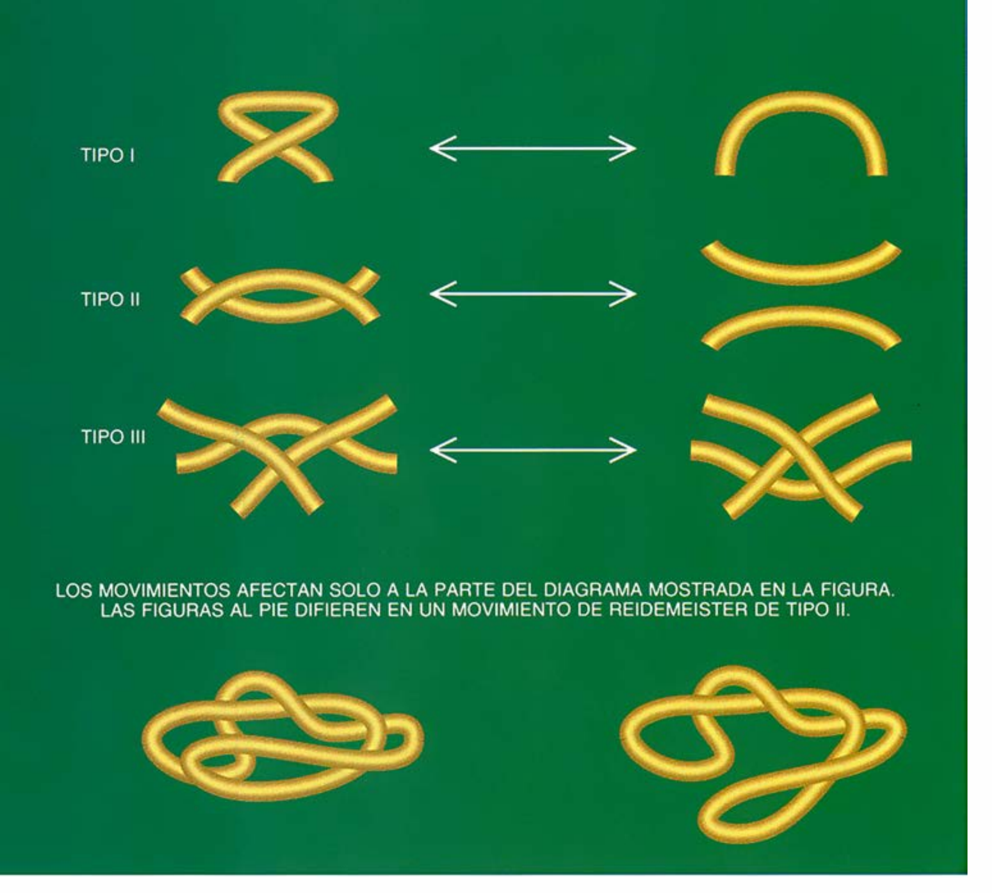
primer lugar, los nudos pueden describirse mediante diagramas bidimensionales (diagramas planares). En segundo lugar, resulta muy difícil distinguir un nudo de otro. No es en modo alguno obvio que dos nudos dados cualesquiera sean diferentes; ni siquiera resulta sencillo saber si un bucle de cordel está anudado o no, pues demostrar tal aserto exige, en principio, considerar la totalidad de las posibles deformaciones del nudo en el espacio tridimensional. La búsqueda de métodos matemáticos que permitan diferenciar a unos nudos de otros, y distinguir los bucles que están anudados de los que no lo están, se ha convertido en uno de los principales problemas de la teoría de nudos.

Kurt Reidemeister, en los años veinte, simplificó bastante el estudio de los nudos introduciendo una pequeña colección de "movimientos" bidimensionales en los diagramas de los nudos. Tales movimientos no cambian el nudo; y dos diagramas cualesquiera del mismo nudo pueden convertirse cada uno en el otro mediante una sucesión de movimientos. Aunque los movimientos de Reidemeister reducen a dos dimensiones el problema de la equivalencia, los movimientos pueden aplicarse de un número infinito de maneras, por lo que el problema no queda en absoluto resuelto.

La más antigua y fecunda técnica de la teoría de nudos deja de lado —al menos en teoría— los diagramas bidimensionales y los movimientos de Reidemeister y se vale, en cambio, de transformaciones topológicas. Dichas técnicas comienzan, por así decirlo, eliminando el nudo del espacio tridimensional ordinario, al objeto de obtener el llamado complemento del nudo; seguidamente, el complemento se somete a deformaciones regulares (continuamente diferenciables, en lenguaje algo más técnico) arbitra-



1. LOS NUDOS Y ENLACES van desde la máxima sencillez a la más enrevesada complejidad. (El nudo que encabeza la figura se basa en una ilustración del *Book of Kells*, libro irlandés del siglo VIII). Los tréboles trifolios levógiro y dextrógiro son las curvas anudadas más simples; antes del descubrimiento del polinomio de Jones, no eran matemáticamente distinguibles, ni tampoco lo eran el nudo llano y el nudo de rizo. El enlace de Hopf constituye la figura más sencilla compuesta por dos bucles encadenados. Aunque los bucles del enlace de Whitehead no están encadenados en sentido estricto, es imposible separarlos. Los aros de Borromeo exhiben también un comportamiento insólito: los tres bucles se hallan encadenados, pero basta cortar sólo uno de ellos para dejar libres a los otros dos.



2. LOS MOVIMIENTOS DE REIDEMEISTER simplifican el análisis de los nudos. Dos nudos son iguales sólo si sus diagramas son idénticos merced a cierta combinación de movimientos.

rias. La estructura topológica del complemento proporciona los llamados invariantes topológicos del nudo [véase "La teoría de nudos", por Lee Neuwirth, en INVESTIGACIÓN Y CIENCIA, agosto de 1979], que son expresiones matemáticas que dependen sólo del nudo, y no de una representación particular.

El invariante más famoso es el polinomio de Alexander, descubierto por el matemático estadounidense James W. Alexander en 1928. El polinomio, denotado $\Delta_K(t)$, se construye teniendo en cuenta los números de los diferentes tipos de cruces que muestra el diagrama del nudo. Para un nudo trifolio, o trébol simple, se tiene $\Delta_K(t) = t - 1 - 1/t$. Versiones diferentemente deformadas del mismo nudo tienen el mismo polinomio de Alexander; los nudos con distinto polinomio son diferentes. Empero, dos nudos pueden tener el mismo Δ sin ser necesariamente equivalentes. El polinomio de Alexander no distingue, por ejemplo, entre el nudo llano y el nudo de rizo [*véase la figura 1*]. Aunque los especialistas han ido desarrollando invariantes adicionales a lo largo de los últimos 60 años, siguen siendo muchos los problemas pendientes de solución.

Era una mañana primaveral y cálida de mayo de 1984 cuando tomé el metro hasta la Universidad de Columbia, para reunirme con Joan S. Birman, especialista en teoría de "trenzas" (que pueden considerarse como un tipo bastante especial de nudos). En mi trabajo sobre álgebras de von Neumann había descubierto, con gran sorpresa, expresiones que guardaban fuerte semejanza con las expresiones algebraicas de ciertas relaciones topológicas entre trenzas. Confiaba en que las técnicas que había estado utilizando pudieran resultar valiosas en teoría de nudos. ¡Quién sabe si podría llegar incluso a deducir nuevas propiedades del polinomio de Alexander!

Retorné a casa un tanto desanimado, tras un largo día de discusiones con Birman. No parecía que mis ideas guardasen relación ni con el polinomio de Alexander ni con ningún otro aspecto de la teoría de nudos.

Pero una noche de la semana siguiente me encontré sentado en mi cama y hube de salir corriendo a efectuar ciertos cálculos. La fortuna me sonrió al usar un planteamiento más sencillo que el que había estado ensayando. Me di cuenta de que había generado un invariante polinómico de los nudos. Lo más verosímil era que se tratase meramente de una nueva expresión del polinomio de Alexander, aunque, por sí misma, tal conexión hubiera sido ya de extremo interés para la mecánica estadística. También podía ser que se tratase de un polinomio nuevo, de un desarrollo inesperado para la teoría de nudos. Era una situación extraña, en la que no se podía perder.

La segunda posibilidad resultó correcta. Llamé $V(t)$ a este nuevo invariante de los enlaces, y no tardé en observar que guardaba semejanza superficial con el polinomio de Alexander. Tanto $V(t)$ como $\Delta(t)$ se pueden calcular de acuerdo con la relación de enmadejado inventada por el matemático británico John Horton Conway.

La semejanza entre $V(t)$ y $\Delta(t)$ espoleó a varios grupos a desarrollar un polinomio de dos variables, hoy conocido por HOMFLY (acrónimo deducido de las iniciales de seis de las ocho personas que lo descubrieron), que da cabida, en particular, a la información contenida en Δ y en V. En este punto, sin embargo, da fin la semejanza entre V y Δ. Todas las tentativas encaminadas a interpretar $V(t)$ en el mismo contexto topológico que $\Delta(t)$ han fracasado. Además, se ignora todavía si existe algún nudo no trivial cuyo Δ sea el mismo que el de un bucle desanudado. En cambio, ya en 1934 se descubrieron nudos cuyo polinomio de Alexander es igual a 1 (como es el caso del polinomio de Alexander de un bucle simple).

En todo lo anterior, no hay indicio alguno acerca de la relación entre la teoría de nudos y la mecánica estadística. La conexión entre ambas teorías no está a la vista; exige cierta explicación de qué es la mecánica estadística y de la clase de problemas para los que es idónea. Lo cual, a su vez, requiere empezar por la mecánica clásica.

En mecánica clásica, un sistema de partículas puede quedar bien determinado dando la posición y el momento lineal (también llamado cantidad de movimiento) de cada una de las partículas en un instante dado. La evolución futura del sistema queda entonces determinada por las leyes físicas. Pero en vista de que un gramo de gas hidrógeno contiene alrededor de 3×10^{23} moléculas, no sería razonable tratar de especificar todas las posiciones y momentos de cada una de las partículas de gas. Además, los cambios producidos en el sistema por eliminación de unas pocas moléculas

CALCULO DE LOS INVARIANTES NODALES

Los invariantes nodales son expresiones matemáticas que describen ciertas propiedades de los nudos. Su nombre se debe a que permanecen invariables cuando se tensa, afloja o retuerce el nudo subyacente. No obstante, la senda conducente desde un bucle de cordel hasta un polinomio en potencias de *t* es harto tortuosa. ¿Cómo se determina un invariante —el polinomio de Alexander o el de Jones, por ejemplo— de un nudo dado? El método más sencillo recurre a la relación de enmadejado, inventada por el matemático británico John Horton Conway. (Siendo alumno de secundaria, Conway había tratado de escribir un programa de ordenador para calcular el polinomio de Alexander, pero halló que la relación de enmadejado funcionaba tan bien que el programa resultaba innecesario.) Otra técnica, el "modelo de estados", de Kauffman, es notable por su elegancia y porque apunta una conexión con la mecánica estadística.

LA RELACION DE ENMADEJADO

La relación de enmadejado parte de tres enlaces orientados (estructuras anudadas y posiblemente formadas por más de un bucle) que son idénticas excepto en la región de un cruce.



L_+



L_-



L_0

En ese cruzamiento, los eslabones se diferencian como vemos. La relación de enmadejado enuncia que

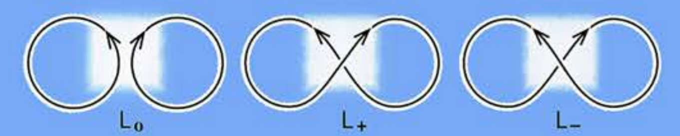
$$(1/t)V_{L+} - tV_{L-} = (\sqrt{t} - 1/\sqrt{t})V_{L0}$$

Además, los V y Δ de un bucle desanudado se normalizan a 1. Estos dos hechos permiten calcular estos dos invariantes para un enlace cualquiera.

El más sencillo de los cálculos no triviales es el V de dos bucles desanudados.



En este enlace hay una sola región de posible cruce; los tres enlaces correspondientes en la relación de enmadejamiento son



L_0 $\quad$ L_+ $\quad$ L_-

L_+ y L_- están ambos desanudados, y sus polinomios son iguales a 1. En este caso, la relación de enmadejamiento se reduce a

$$1/t - t = (\sqrt{t} - 1/\sqrt{t})V_{L0}$$

Dividiendo ambos miembros entre $(\sqrt{t} - 1/\sqrt{t})$ resulta

$$V_{L0} = -(\sqrt{t} + 1/\sqrt{t})$$

El proceso es un poco más difícil en el caso de un enlace más complejo, como el trébol. Se construye, como antes, un sistema de enlaces que se diferencien únicamente en un cruce.



L_+ (el trébol)



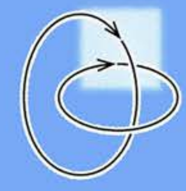
L_- (desanudado)



L_0 (un enlace de Hopf)

La relación de enmadejado nos da V_L en función de V_H:

Para calcular V_H, se construye otro sistema de enlaces:



L_+ (un enlace de Hopf)



L_- (dos bucles)



L_0 (desanudados)

Aplicando la relación de enmadejamiento a este sistema resulta

$$(1/t)V_H + t(\sqrt{t} + 1/\sqrt{t}) = (\sqrt{t} - 1/\sqrt{t})$$

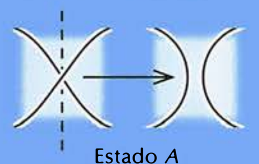
y en definitiva $\quad V_H = -\sqrt{t}\,(1 + t^2)$

Al sustituir este valor en la relación de enmadejamiento del trébol resulta

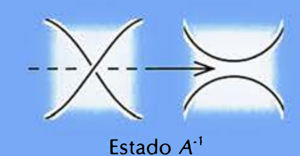
$$V_L = t + t^3 - t^4$$

EL MODELO DE ESTADOS

El modelo de estados de Kauffman opera eliminando todos los cruces de un enlace y reemplazando el enlace por círculos desconectados. Cada cruce puede encontrarse en uno de dos estados, que determina cómo será eliminado; cada posible combinación de estados de cruce determina un estado del enlace.



Estado *A*



Estado A^{-1}

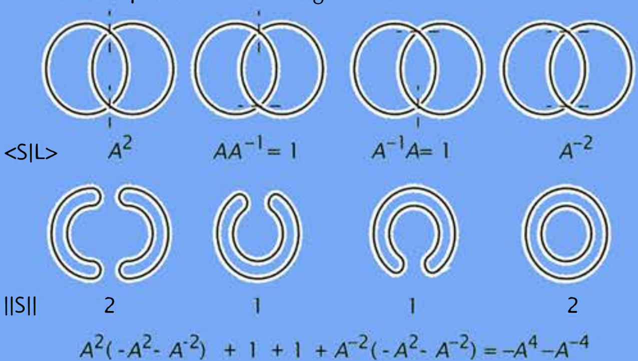
Kauffman define entonces el "polinomio corchete" correspondiente a un enlace mediante una suma extendida a todos los estados del enlace.

$$\langle L\rangle = \sum_S \langle S|L\rangle\,(-A^2 - A^{-2})^{\|S\|-1}$$

($\langle S|L\rangle$ denota el producto de todas las *A* y 1/*A* en los cruces, y $\|S\|$ es el número de círculos del diagrama resultante.)

Por ejemplo, el corchete del enlace de Hopf es $(-A^4 - A^{-4})$.

Los estados posibles son los siguientes



$\langle S\|L\rangle$	A^2	$AA^{-1} = 1$	$A^{-1}A = 1$	A^{-2}
$\|S\|$	2	1	1	2

$$A^2(-A^2 - A^{-2}) + 1 + 1 + A^{-2}(-A^2 - A^{-2}) = -A^4 - A^{-4}$$

Para obtener V(t) a partir del polinomio corchete basta hacer $A^4 = t$ y multiplicar el corchete por una potencia adecuada de A:

$$-A^4 - A^{-4} = A^{-6}(-A^2(1 + A^8)) = A^{-6}(-\sqrt{t}\,(1 + t^2)) = A^{-6}V_H(t)$$

le resultaría completamente inapreciable a quienquiera que estuviese observando el sistema completo.

Las únicas magnitudes que revisten interés para la mecánica estadística son las que se muestran insensibles a variaciones microscópicas; por ejemplo, la energía media (la temperatura) de un conjunto de moléculas. Si imaginamos un gran sistema construido por aportación de átomos de uno en uno, dichas magnitudes tenderán hacia un límite bien definido al tender hacia infinito el tamaño del sistema.

Por ingenuo que pueda sonar, el examen de la conducta de los agregados de partículas da lugar a paradojas. Una de las más evidentes es la irreversibilidad de muchos procesos. Las leyes de movimiento no varían al invertir el sentido de progresión del tiempo. Una colisión elástica entre una bola y un obstáculo, sea por caso, ofrece idéntico aspecto haciendo correr el tiempo hacia adelante o hacia atrás. Mas pensemos ahora en un sistema de bolas que chocan y rebotan sin fricción en una mesa rectangular, limitadas por una pared a moverse en una mitad de la mesa. Al retirar la pared, las bolas se dispersarán rápidamente por todo el tablero; jamás se pondrán de acuerdo para retornar a la mitad de la que salieron. El mero hecho de que un sistema contenga un gran número de partículas parece dotar al tiempo de un sentido definido.

CONEXION DE LA TEORIA DE NUDOS CON LA MECANICA ESTADISTICA

La conexión entre estos dos dominios de apariencia tan dispar, uno de los cuales trata de grandes agregaciones de partículas de comportamiento individual muy simple, mientras el otro se ocupa sobre todo de sistemas pequeños que exhiben sutiles rasgos, resulta más fácil de señalar desde la mecánica estadística. Para remedar fenómenos naturales como la fusión o la ebullición, los físicos han recurrido a modelos espinoriales: agregados simplificados definidos sobre retículos de puntos. Cada punto posee un "espín" particular e interactúa con su vecino más cercano, de acuerdo con una función de energía que depende de los dos espines. La "función de partición" de un modelo es

$$\sum_{\text{ESTADO}} e^{-E(\text{ESTADO})/kT} \qquad (k \text{ es la constante de Boltzmann})$$

La relación estrella-triángulo, que solamente se verifica para ciertas clases de modelos espinoriales (entre los que se cuentan el modelo de Ising y el modelo de Potts del Q-estado), asevera que las funciones de partición de los modelos espinoriales definidos sobre los dos grafos siguientes son proporcionales.

A B C

A B D C

$$e^{-(E(A,B)+E(A,C)+E(B,C))/kT} = R\sum_{D} e^{-(E(A,D)+E(B,D)+E(C,D))/kT}$$

La relación anterior se verifica también para los llamados modelos-vértice, en los que asignan espines a cada lado del grafo y en los que la energía se calcula de acuerdo con los cuatro espines que concurren en un vértice. La relación estrella-triángulo para un modelo de vértice se expresa mediante la ecuación de Yang-Baxter:

T Z Y S A C B X R

T Z B A C S Y X R

$$\sum_{BSY} e^{-\left\{\frac{E(A,B|X,Y)+E(R,S|B,C)+E(Y,Z|S,T)}{kT}\right\}} = \sum_{BSY} e^{-\left\{\frac{E(A,B|S,T)+E(X,Y|R,S)+E(B,C|Y,Z)}{kT}\right\}}$$

Otra de las sorpresas que ofrecen los sistemas grandes es la existencia de transiciones de fase. El hielo se funde; el agua hierve. Ciertas magnitudes, como la presión, que para sistemas pequeños se pueden definir mediante funciones regulares, exhiben discontinuidades espontáneas. Basta considerar sistemas que contengan muchas partículas para convertir en discontinuas a tales funciones.

Las transiciones de fase manifiestan un interesante comportamiento cualitativo; por ello, los físicos han desarrollado sistemas definidos abstractamente —llamados "modelos"— en los cuales resulta posible calcular explícitamente magnitudes macroscópicas, como la presión o la capacidad calorífica, en función de otros parámetros, por ejemplo, la temperatura.

Pero estas magnitudes macroscópicas resultan muy difíciles de calcular incluso en casos sencillos. En consecuencia, los modelos tienden a evitar un realismo minucioso. El modelo más sencillo es el modelo de Ising, resoluble en dos dimensiones, que consiste en un sistema de "espines" dispuestos en una configuración regular. (Los espines son magnitudes puramente matemáticas; no es forzoso que posean significado físico intrínseco.) Cada espín interactúa exclusivamente con sus vecinos más próximos; dando a cada uno de los espines un valor de +1 o −1 queda definido un "estado" del sistema. La energía de un estado viene dada por la suma de las energías resultantes de las interacciones entre vecinos inmediatos.

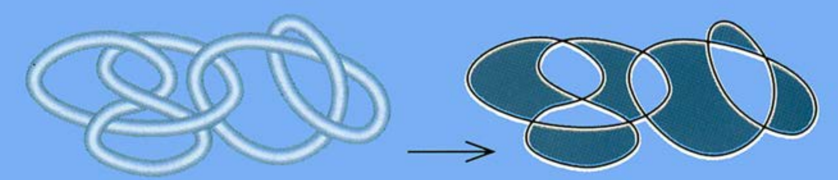
Muchas magnitudes significativas para sistemas grandes resultan deducibles de la llamada función de partición, o función Z, que se define como la suma, extendida a toda la colección de estados, de las exponenciales de las energías cambiadas de signo. El modelo de Ising puede tener un número enorme de estados; por ello, en el cálculo de esta suma aparecen todos los problemas de la resolución de un modelo. Incluso informatizando los cálculos, el análisis tiene que reducirse a retículos pequeños. Sin embargo, se han ideado diversas e ingeniosas técnicas matemáticas. El matemático noruego Lars Onsager resolvió, en 1944, el modelo de Ising con retículo bidimensional cuadriculado; su fórmula, que abrió el camino hacia nuevos avances, mostraba la existencia de una transición de fase en el modelo.

Uno de los resultados importantes del trabajo de Onsager fue el conocido por relación estrella-triángulo. La relación estrella-triángulo correspondiente a un modelo de Ising definido sobre un par concreto de grafos enuncia que las funciones de partición han de ser necesariamente múltiplos escalares una de otra en cuanto se fijan en los lados de los grafos los valores de los espines. La relación debe su nombre al hecho de que uno de los grafos es triangular y, el otro, asteriforme.

Rodney J. Baxter, de la Universidad Nacional Australiana, ha elaborado una solución del modelo de Ising enteramente fundada en las consecuencias de esta relación. Una mara-

DE NUDOS A GRAFOS

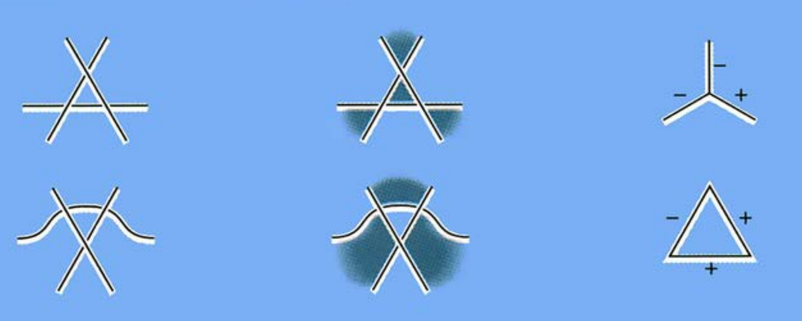
El tránsito de los modelos espinoriales y la mecánica estadística a la teoría de nudos exige la conversión de los nudos en grafos. Se empieza coloreando el nodo (o el enlace) en dos tonos, a modo de damero:



Para extraer un grafo, tomemos como vértices las regiones oscuras y, como lados, los cruces entre ellas. Los lados se rotulan como positivos o negativos de acuerdo con la orientación relativa del cruzamiento y de las regiones oscuras:



Un movimiento de Reidemeister de tipo III aplicado al grafo de un nudo produce inmediatamente la relación estrella-triángulo:



villosa peculiaridad del método de Baxter es la de prestarse a generalizaciones. El modelo de Ising puede generalizarse permitiendo que los espines de una ubicación tomen valores distintos de $+1$ y -1. Tales modelos se denominan modelos espinoriales, y están definidos por el conjunto de energías correspondientes a interacciones entre ubicaciones vecinas para todos los valores del espín.

En general, los modelos espinoriales no satisfacen la relación estrella-triángulo. Para aplicar el trabajo de Baxter, es necesario hallar conjuntos de energías de interacción que verifiquen las ecuaciones triángulo-estrella, y resolver los modelos definidos por esos conjuntos de energías. La solución más sencilla está constituida por el modelo de Potts del Q-estado autodual. Este modelo admite para el espín Q valores posibles (a diferencia de los dos que permite el modelo de Ising); y dos vecinos interactúan tan sólo cuando sus espines son iguales. (Los primeros en resolver este modelo, si bien por otro método, fueron Neville Temperley, a la sazón en la Universidad de Swansea-Gales, y Elliot H. Lieb, de la Universidad de Princeton.)

¿Qué relación puede existir entre la teoría de nudos y los modelos de la mecánica estadística? El primer paso para establecer la conexión consiste en sombrear alternativamente de dos tonos las regiones de un diagrama de enlace (el dibujo de un nudo), como si fuesen los escaques de un damero. De este "damero" se puede extraer un grafo tomando como vértices las regiones oscuras y tomando los cruces como lados. Seguidamente, se procede a atribuir signos a los lados de acuerdo con la orientación relativa entre el cruce y la configuración del damero. (El recuadro expone el procedimiento completo.) ¡Un movimiento Reidemeister de tipo III produce inmediatamente la relación estrella-triángulo!

El movimiento que deja invariable al nudo deja también esencialmente invariable a la función de partición de un modelo de Ising (sea por caso) derivado del diagrama de enlace. Tal observación hace del todo natural el examen de las funciones de partición correspondientes a modelos espinoriales definidos sobre el grafo deducido del sombreado bicolor de un enlace, con el fin de averiguar si también ellas describen invariantes de ese nudo.

La función de partición depende sólo de la suma, extendida a todos los estados, de la función de energía del modelo espinorial. Es decir, cada estado define los espines de todos los vértices del grafo y, por tanto, la energía de las interacciones entre ellos. Para calcular la función de partición, se asigna a cada lado del grafo la exponencial de la energía de interacción de los vértices que ese lado conecta. El proceso se repite después para la totalidad de estados posibles. (Para tener en cuenta los signos de los cruces, es preciso postular dos funciones de energía para la interacción entre vértices adyacentes, una positiva y una negativa.)

La cuestión que ahora se plantea es: ¿qué condiciones han de satisfacer las dos funciones de energía para constituir un invariante del enlace? (Ello significa que el valor de la función de partición no debe cambiar si se aplica un movimiento de Reidemeister al diagrama subyacente.) La mejor solución consiste en retornar a la mecánica estadística, pues la función de energía de los modelos que satisfacen la relación estrella-triángulo (y demás relaciones correspondientes a movimientos de Reidemeister) engendrará funciones de partición que son también invariantes de enlace. El modelo de Ising, por ejemplo, engendra un invariante conocido en teoría de nudos por invariante Arf o invariante de Kervaire.

Y lo que reviste mayor importancia: el modelo de Potts del Q-estado produce un invariante para cada número Q de valores admitidos para el espín. A decir verdad, tal invariante es el polinomio $V(t)$. Q y t están relacionados por la fórmula $Q = 2 + t + t^{-1}$ (relación fácil de verificar comprobando la relación de enmadejado).

El truco del sombreado bicolor ha establecido una auténtica conexión entre la mecánica estadística y la teoría de nudos, si bien por un método francamente extravagante. La relación estrella-triángulo apareció en la mecánica estadística como método para la resolución de modelos. Apareció en teoría de nudos como requisito para la invariancia topológica. Si no fuera por la ubicuidad con que interviene, podría pensarse que una y

otra razón son del todo independientes; que la conexión entre ambas es meramente fortuita.

Si en lugar de los modelos de Potts e Ising se utilizan otros modelos espinoriales, aparecen otros invariantes para nudos, también utilizados por los matemáticos. Existen además otros tipos de modelos que son más potentes aún desde el punto de vista de los invariantes de los nudos. En los modelos de vértices, por ejemplo, los lados tienen asignados espines mientras que las energías provienen de los vértices. Merced a los modelos de vértices y a ciertos grupos felizmente llamados "grupos cuánticos" (que son sistemas de simetrías más refinadas que las meras simetrías geométricas del espacio ordinario), resulta posible construir una pléyade de invariantes polinómicos; tantos, que bastan para determinar completamente el polinomio HOMFLY de dos variables y otro polinomio de dos variables conocido por polinomio de Kauffman.

Así pues, aunque resulte misteriosa, existen indicios sustanciales de la existencia de una conexión entre la teoría de nudos y la mecánica estadística. Muchas de las esperanzas de lograr una comprensión más profunda de esta correspondencia se centran en la teoría de campos, una visión matemática del mundo en la que cada punto del espacio es portador de una variable. (Por ejemplo, en un fluido en movimiento, la velocidad del fluido se describe mediante un campo vectorial; cada punto tiene asociado un vector que expresa la velocidad.)

Un método muy potente para abordar la teoría de campos consiste en construir el campo como límite de campos definidos sobre retículos discretos cuya amplitud de malla tiende a cero. De esta forma, los modelos de la mecánica estadística, como el modelo de Ising, tienen aplicación también en teoría de campos. De hecho, una inspección minuciosa de los procesos utilizados muestra que el límite continuo del modelo de Ising bidimensional es una teoría cuántica de campo unidimensional. (Las teorías cuánticas de campo son teorías cuyas variables son operadores que actúan sobre estados cuánticos.)

Algunos matemáticos y físicos han tratado de explicar la presencia de la teoría de nudos en los modelos mecánico-estadísticos examinando su comportamiento en el límite continuo. Aunque sus esperanzas de hallar una solución no se han visto cumplidas todavía, Witten ha puesto de manifiesto que el lenguaje de la teoría cuántica de campos facilita al menos un armazón o referencia sobre el que definir los invariantes descritos en este artículo.

Resulta revelador que el trabajo de Witten demuestre la forzosa existencia de los invariantes en cualquier espacio tridimensional. Además del espacio tridimensional en que vivimos, hay muchos otros, genéricamente denominados variedades tridimensionales [véase "Matemática de las variedades tridimensionales", por William P. Thurston y Jeffrey R. Weeks, INVESTIGACIÓN Y CIENCIA, septiembre de 1984]. El toro tridimensional, por ejemplo, se construye estirando un cubo y "pegando" sus caras superior e inferior, izquierda y derecha, delantera y trasera, con lo cual una partícula que saliera del cubo hacia la izquierda volvería en el acto a introducirse en él desde la derecha. (No es posible efectuar materialmente esta operación en el espacio tridimensional.)

Más todavía. La teoría de Witten asevera la existencia de invariantes para los enlaces contenidos en una variedad tridimensional arbitraria, incluso para enlaces sin componentes. Se sigue que los invariantes suministran información topológica no sólo de los meros enlaces, sino también acerca de las propias variedades tridimensionales. Son muchos los matemáticos y físicos que están investigando activamente este nuevo desarrollo, que pudiera forjar conexiones entre la teoría de nudos y otros campos de la ciencia a primera vista menos relacionados aún con ella que lo estaba la mecánica estadística de hace seis años.

BIBLIOGRAFIA COMPLEMENTARIA

EXACTLY SOLVED MODELS IN STATISTICAL MECHANICS. R. J. Baxter. Academic Press, 1982.

A POLYNOMIAL INVARIANT FOR KNOTS VIA VON NEUMANN ALGEBRAS. V. F. R. Jones en *Bulletins of the American Mathematical Society*, vol. 12, Parte 6, págs. 103-111; 1985.

POLYNOMIALS FOR LINKS. W. B. R. Lickorish en *Bulletins of the London Mathematical Society*, vol. 20, págs. 558-588; 1988.

QUANTUM FIELD THEORY AND THE JONES POLYNOMIAL. Edward Witten en *Communications in Mathematical Physics*, vol. 121, n.º 3, págs. 351-399; 1989.

REPRESENTATIONS OF THE ALGEBRA $U_q(SL_2)$, Q-ORTHOGONAL POLYNOMIALS AND INVARIANTS OF LINKS. A. N. Kirillov y N. Yu. Reshetikhin en *New Developments in the Theory of Knots*. Dirigido por T. Kohno. World Scientific Publishing, 1990.

Una fábrica romana

La mayoría de los historiadores sostienen que los romanos no desarrollaron la técnica de la producción mecanizada por lo barato que era el trabajo de los esclavos. Las ruinas de Barbégal, en Francia, cuentan una versión distinta

Trevor Hodge

Pocos visitantes de Arles, en el sur de Francia, suelen tomarse la molestia de prolongar su viaje hasta la aldea de Fontvieille, sita a 19 kilómetros al nordeste. Menos aún son los que giran hacia el sur desde Fontvieille y siguen cuatro kilómetros por la carretera D 33, donde un camino de carro, que se desvía a la izquierda, lleva hasta una de las maravillas técnicas de la antigüedad clásica, el molino de Barbégal.

Barbégal es un ejemplo bien conservado de algo que, a tenor de los libros de texto, jamás ha existido: un conjunto fabril auténtico, de la época romana, impulsado por energía y dedicado a la producción en masa. Es un yacimiento grande, imponente, que no requirió excavación alguna. Pero la atención concentrada en templos y otros monumentos artísticos no se detuvo en los monumentos técnicos como los molinos; hizo falta que el arqueólogo francés Fernand Benoît publicase sus investigaciones en 1940, para que la ciencia se percatase de su importancia.

Barbégal encierra su interés porque pone en tela de juicio lo que se puede denominar teoría tecnológica de la decadencia y caída del Imperio Romano. Esta teoría sostiene que la facilidad de encontrar mano de obra barata –los esclavos– disuadió a los romanos de poner en marcha fuentes de energía alternativas sin las cuales es imposible la producción de manufacturas en gran escala.

A. TREVOR HODGE es profesor de estudios clásicos en la Universidad de Carleton, Ottawa. Estudió en Cambridge, donde se doctoró en arqueología, se dedicó en un principio a la arquitectura griega, en particular a las técnicas de construcción de templos, pero después ensanchó el círculo de sus intereses para abarcar la campaña de la Batalla de Maratón y la fundación de Marsella por los griegos.

No hace falta más que recorrer unos cien metros camino abajo, para ver los restos de este complejo industrial, que desciende asido a la ladera. Su explicación es sencilla. Barbégal era un inmenso molino harinero, que data probablemente del siglo IV d.C. La energía que accionaba las ruedas de molino procedía de 16 norias colocadas en filas paralelas de ocho. Cada fila descendía cuesta abajo, de modo que el agua se precipitaba de cada rueda a la siguiente, haciendo girar las ocho antes de perderse en el desagüe situado al pie de la cuesta, que desembocaba en una ciénaga, distante 500 metros.

El ingenio del dispositivo salta a la vista. En la época medieval y posterior no era raro construir molinos uno tras otro siguiendo el curso de una misma corriente; aquí estaban conjuntados bajo un mismo techo para formar una verdadera fábrica. También es admirable el mecanismo desde el enfoque opuesto, pues las fuentes de energía estaban distribuidas inteligentemente. Hubiera sido imposible accionar tantas máquinas con una sola noria porque la fricción de la correa impulsora, o de cualquier otro medio de transmisión, hubiese constituido una carga demasiado grande. Las muelas múltiples eran únicamente posibles si cada una estaba provista de su propia noria.

¿Quién fue el gran ingeniero de Barbégal, título que le corresponde con toda propiedad? Una inscripción sepulcral de fecha plausible, conservada aún en el cementerio Aliscamps de Arles, rinde homenaje a un cierto Quintus Candidius Benignus, famoso ingeniero hidráulico de la localidad, que era "diestro como ningún otro, y nadie le superaba en la construcción de máquinas y en el tendido de redes de canalización pública". No hay una prueba segura de que fuese él quien construyó Barbégal o de que la fábrica fuese la razón de su fama, pero a uno le gustaría pensar que así fue.

Subiendo por la cuesta hasta la cima, se ve al otro lado una vaguada, cruzada por un acueducto típicamente romano que abastecía el molino. Observándolo más detenidamente, se advierte que otro acueducto se alza al lado del primero, que lo oculta en gran parte. En la divisoria, el acueducto del lado oeste giraba repentinamente a la derecha y continuaba hacia el oeste, en una prolongación evidente de 17 kilómetros para abastecer de agua potable a Arelate, la ciudad romana de Arles. El acueducto oriental continuaba en línea recta, atravesando la cresta por un canal especialmente excavado, hasta alcanzar el extremo superior del complejo de molinos que tenía que abastecer.

El molino disponía de un acueducto para su uso exclusivo; eso es obvio. Se discute, sin embargo, la relación precisa entre el acueducto del molino y el acueducto de la ciudad. Los últimos estudios parecen indicar que comenzó por construirse un acueducto para abastecer la ciudad, y que después de un considerable lapso de tiempo quedó inservible, ya fuera porque la cantidad de agua disminuyese, ya porque la conducción quedase atascada por la incrustación de carbonato de calcio que sufrían la mayoría de los acueductos. En esta tesitura, sigue la argumentación, se construyó un nuevo acueducto; el original pasó a convertirse en una simple fuente de energía. Por consiguiente, el molino se habría construido para explotar una fuente de energía inesperada en un emplazamiento apropiado, pues tanto la calidad como la cantidad del agua todavía eran adecuadas para un empleo industrial.

Al salir del canal en lo alto del cerro, el acueducto se bifurcaba en forma de Y. Se creyó tiempo atrás que el espacio triangular delimitado por la Y se usaba como depósito de cabecera (para asegurar un abastecimiento regular de agua), pero esta hipótesis ahora se considera improbable. El complejo molinero propiamente

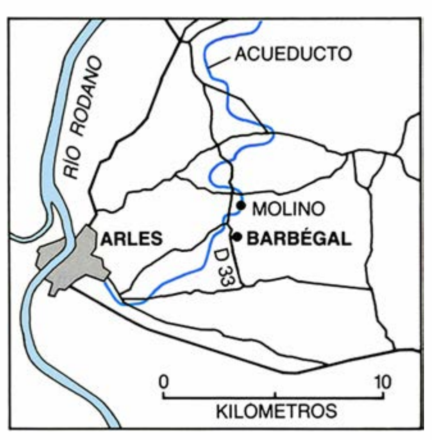
dicho, en que el agua entraba a partir de aquí, estaba dispuesto bajando la cuesta de la ladera formando un rectángulo de aproximadamente 42 por 20 metros. Se componía de dos filas de molinos, donde el grano era triturado, separadas por una vía de servicio que permitía a mulos y asnos acarrear el grano y recoger los sacos de harina. Por la parte de fuera de las paredes iban montadas dos filas de ruedas hidráulicas, cada una de ellas encerrada dentro de un recinto oblongo de cantería llamado caja de rueda. Estas cajas disponíanse, a su vez, una tras otra en una fila escalonada que bajaba por la pendiente. El espaciado entre los emplazamientos de los molinos, en relación con el de las ruedas instaladas en el exterior de la pared común, no era del todo uniforme, según queda patente. Las cajas de algunas ruedas son más largas que otras, y la parte baja de la pared no es del todo recta. Estas discordancias venían impuestas sin duda por desigualdades del terreno.

Otras irregularidades se derivan, por lo menos en parte, de la dificultad de hacer una reconstrucción completa. En ningún sitio quedan en pie paredes de más de un metro y medio de altura. La mayor parte de su cuerpo alto era de madera, desaparecida tiempo ha. El restaurador se ve obligado, por tanto, a desplegar una buena dosis de imaginación creadora, a la par que rigurosa. Esta circunstancia no es rara en la restauración de edificios antiguos. Aquí, sin embargo, el investigador se enfrenta con un problema añadido y puramente técnico; a saber: el descenso de nivel de la colina, desde la cima hasta el pie, tiene una pendiente de 30 grados por término medio; mas, para acoplar las ocho ruedas hidráulicas, el curso del agua habría de tener una pendiente bastante más pronunciada. Aquí son varias las soluciones propuestas; una de ellas, la de levantar la parte superior del molino hasta un nivel alto y rebajar el tramo inferior, instalando las ruedas hidráulicas en los cimientos, con lo que la pendiente hidráulica quedaría acentuada. Aunque esta di-



1. CIMENTACION EN RUINAS de un molino harinero romano que sube por una ladera en Barbégal. Un acueducto alimentaba dos corrientes de agua que caían sobre 16 ruedas hidráulicas, cuyos emplazamientos en caja aún se distinguen. Es la mayor fábrica de la Antigüedad que se conoce.





2. BARBEGAL fue construido en el siglo IV d.C. cerca del puerto de Arles (*izquierda*), al lado de un acueducto que en tiempos había abastecido a la ciudad (*derecha*). El molino producía probablemente harina suficiente para alimentar a las 12.500 personas que vivían en Arles en aquel entonces.

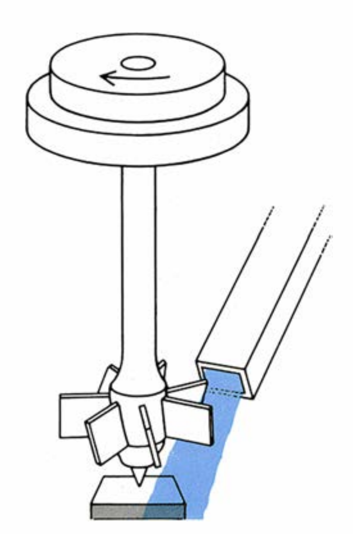
ferencia en el emplazamiento de las ruedas pudiera tener un reflejo arquitectónico en la disposición sensiblemente distinta de las tres últimas casetas, nadie ha podido demostrar cómo solucionaron los romanos la disparidad entre la pendiente hidráulica y la topográfica.

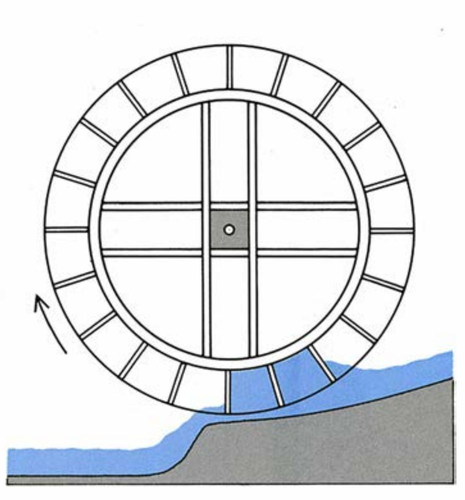
Cada rueda medía 0,7 metros de ancho por 2,1 de diámetro. Iba montada en un eje que pasaba desde su caja, a través de un agujero de la pared (todavía se distinguen uno o un par de ellos), hacia el interior de la caseta del molino. Allí se conectaría, mediante un par de ruedas dentadas, a un eje vertical que subía y atravesaba el techo; en el piso de arriba impulsaba la piedra giratoria del molino (una pieza de basalto de 90 centímetros de diámetro) que era la que molía. El sistema es bien conocido gracias a otros molinos romanos. La principal laguna en nuestros conocimientos dimana de que no sabemos si se sacaba algún provecho del sistema de transmisión para incrementar la velocidad de rotación, y con ello el rendimiento de la molienda.

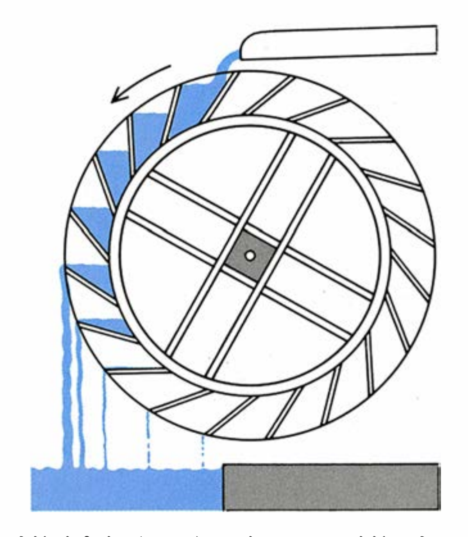
Este eslabón perdido ilumina lo que constituye quizás el enigma más angustioso de Barbégal: ¿Qué potencia generaba el molino y cuánta harina podía producir, o de hecho producía? Las dos últimas preguntas no son necesariamente idénticas, porque no sabemos si el molino funcionaba continuamente o cerraba por la noche. Además, el tamaño de las piedras de moler lo conocemos, pero no su velocidad de revolución. Pese a ello, la cifra aproximada de la producción puede dar una idea de la magnitud en cuestión.

El cálculo más reciente fue realizado en 1983 por Rober H.J. Sellin, basándose en el caudal de un acueducto de un metro por segundo (descargando 0,3 metros cúbicos por segundo) y en una rueda que girase diez veces por minuto al 65 por ciento de rendimiento. Sellin calcula que una rueda de ese tipo podría generar unos dos kilowatts, o sea, 2,5 caballos. El comentario del estudioso británico John Landels es de este cariz: “Para hacer una comparación moderna (algo deprimente), el motor de una motocicleta pequeña (250 centímetros cúbicos) desarrolla aproximadamente la misma potencia”. Esta potencia, con la transmisión adecuada, haría girar una piedra de molino a unas 30 revoluciones por minuto, suficiente para producir unos 24 kilogramos de harina molida en 24 horas, o nueve toneladas métricas al día en la totalidad de la instalación de Barbégal. Suponiendo un 50 por ciento de tiempo perdido, para dar margen a interrupciones, retrasos en la llegada de las remesas de grano, mantenimiento, descenso de caudal, averías, tiempo de descanso y otras interferencias, Barbégal aún produciría, según Sellin, “4,5 toneladas métricas diarias, que bastan para alimentar a una población de 12.500 personas tomando como unidad de consumo 350 gramos por día. Esto se ajusta a los cálculos de la población de Arles en el siglo IV d.C.”.

El nuevo cálculo parece resolver un problema. Los anteriores habían sido más elevados, y la mayoría de los historiadores andaban desconcertados ante la posibilidad de que el molino produjese más harina de la que la población local consumía. La explicación socorrida decía que en Arles se había acantonado una legión y que Barbégal abastecía a la guarnición. Pero las guarniciones legionarias no constituían ninguna rareza, y en ninguna otra parte del Imperio se sabe de una guarnición que haya recurrido a semejante procedimiento. El cálculo más moderado elimina estas di-







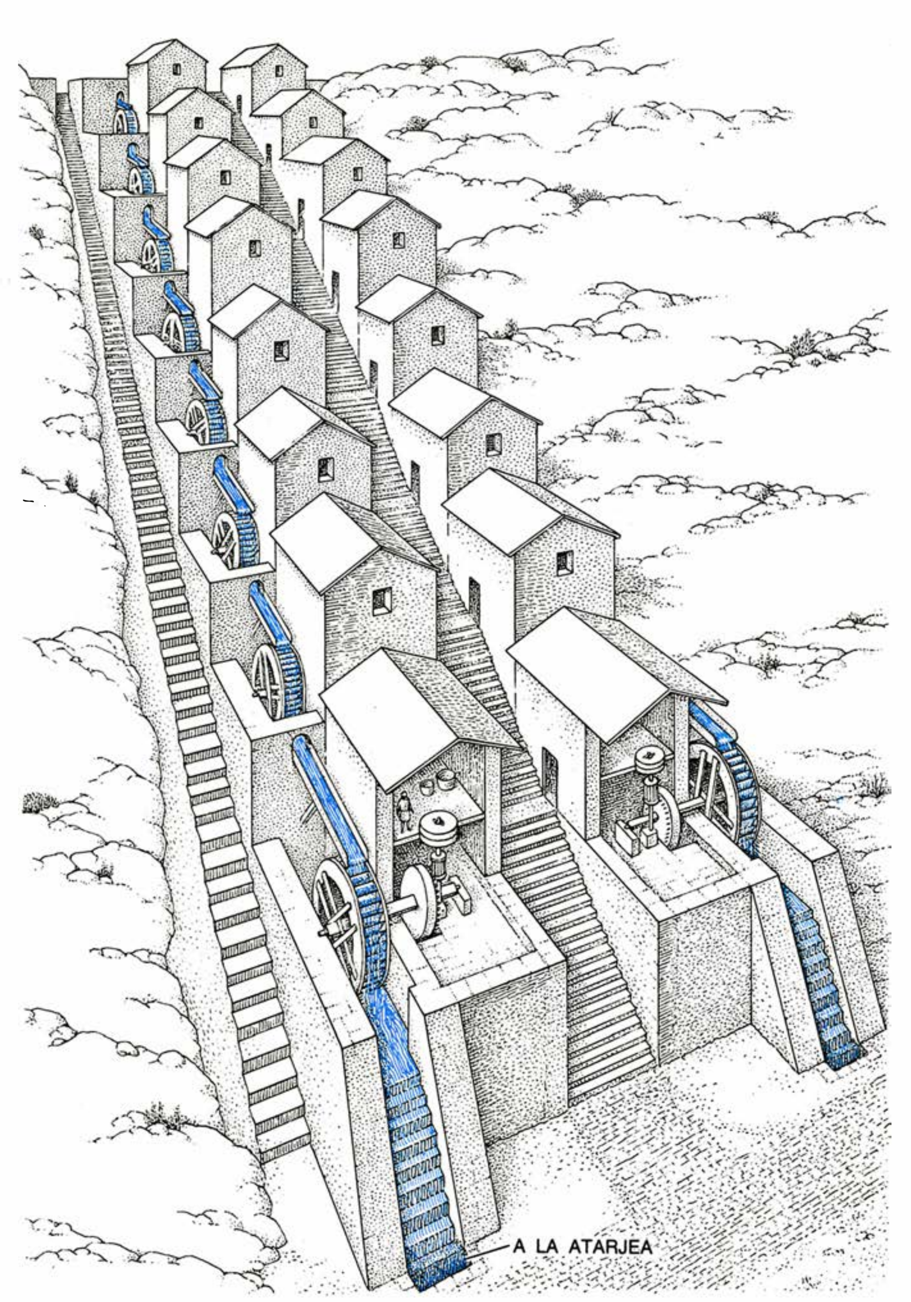
3. RUEDAS HIDRAULICAS en orden de complejidad y productividad crecientes. Las ruedas escandinavas (*izquierda*) hacen girar las muelas directamente, las de propulsión inferior (*centro*) requieren transmisión y las de propulsión superior (*derecha*) exigen una corriente de agua elevada.

ficultades y apoya así la teoría de que la fábrica se construyó para aprovechar el acueducto.

Al determinar el diseño de la fábrica se plantea otro problema técnico que atañe a la clase de rueda hidráulica empleada. En la Antigüedad se conocían tres variantes. La más sencilla –la rueda escandinava– está montada horizontalmente en un eje vertical que atraviesa el suelo del molino haciendo girar la muela de encima. La ausencia de transmisión facilita la construcción y el mantenimiento, pero comporta que la volandera gire lentamente, a la velocidad de la rueda hidráulica, y muele, por tanto, con poco rendimiento. Las ruedas escandinavas no están inmersas en el río, sino accionadas por el chorro de un caño. Todavía son corrientes en China (quizá tenga que ver con esto que allí los molinos de viento también giren en un plano horizontal).

La rueda de agua más corriente, la vertical, ofrece dos variantes, de propulsión inferior y superior. La primera rueda está sumergida en la corriente de agua que la hace girar. La segunda es accionada por el agua de un caz que vierte agua en lo alto de la rueda, impulsándola tanto por el momento del flujo como por la gravedad del agua al caer. En ambos casos, el impulso es recogido por un eje horizontal y transmitido por mecanismos de madera, como ocurría en Barbégal, a un eje vertical. De las dos versiones, la de mayor rendimiento es la de propulsión superior. Pero la eficiencia energética puede perder atractivo si se cuenta con agua abundante para mover el molino sin problema. Además, una rueda impulsada por arriba requiere que el suministro de agua esté regulado por una presa y conducido por un caz elevado; en suma, una instalación muy costosa.

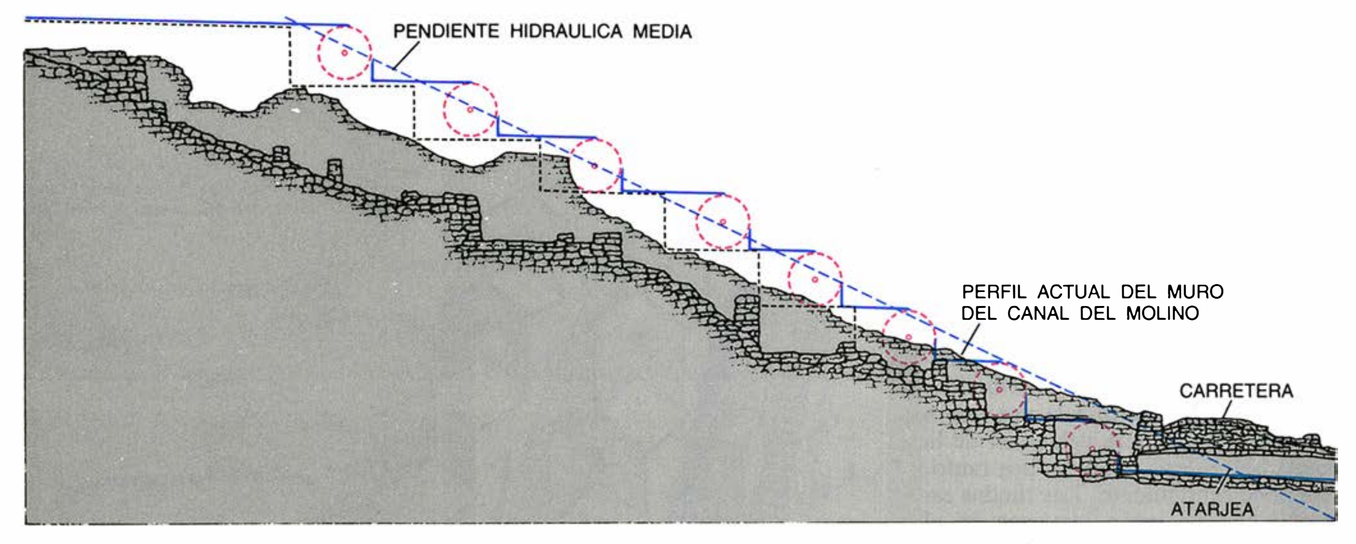
Como hay gradiente paulatino de calidad entre el modelo escandinavo, el propulsado por abajo y el propulsado por arriba, el supuesto tradicional era que las diferentes técnicas habían aparecido en este mismo orden cronológico. Pero N.A.F. Smith ha demostrado que se desarrollaron de manera independiente, en respuesta a necesidades y tradiciones regionales, y que en la mayor parte del Imperio la rueda horizontal era rara, si es que se usaba. En Barbégal es evidente que las ruedas eran verticales. La única duda estriba en saber si eran de propulsión superior o inferior. Aunque hay opiniones en favor de una y otra, no cabe duda de que el primero de estos modelos es el más probable, aunque sólo fuese porque se requiriera su mayor rendimiento para explotar el volumen, pequeño aunque controlado, del agua suministrada por el acueducto.



4. RUEDAS HIDRAULICAS propulsadas por arriba, el modelo más eficaz de los tres principales. Se supone que movieron los molinos (mostrados aquí en una reconstrucción idealizada). Tres escaleras de servicio flanqueaban los molinos, cada uno de ellos probablemente con su propio recinto.

Barbégal trasciende estos aspectos técnicos para incardinar la cuestión del papel social de la técnica. La doctrina ortodoxa sostiene que los antiguos conocían perfectamente la rueda hidráulica, pero hacían de ella un uso sorprendente por su parquedad, debido a que la fuerza motriz proporcionada por los esclavos era sumamente barata.

Aquí la rueda hidráulica, trivial en apariencia, incide en una de las grandes paradojas de la Antigüedad: el relativo estancamiento de la técnica en una sociedad que produjo algunas de las mejores cabezas de la historia. Se acepta de ordinario que la técnica no sacaba el debido rendimiento a la energía aplicada. Ni griegos ni romanos no podían, por ejemplo, aprovechar el vapor porque carecían de los conocimientos para trabajar el metal con maquinaria apropiada. ¿Por qué, pues, despreciaron las fuentes que tenían a su alcance? Los antiguos estaba familiarizados, merced a las velas de los barcos, con la energía del viento, pero al parecer nunca la emplearon en molinos. Construyeron ruedas hidráulicas, pero no las suficientes.



5. LA RECONSTRUCCION de Barbégal emplaza las ruedas a lo largo de una curva descendente, el medio lógico de reconciliar las pendientes del terreno y del agua. Las ruedas estaban probablemente instaladas por encima del suelo en lo alto del cerro y, en el subsuelo, al pie de la ladera.

Llegados a este punto, el estudio reciente ha puesto de nuevo en tela de juicio la versión que se daba por segura. Es muy probable que las ruedas hidráulicas no fuesen tan raras en la Antigüedad. Los testimonios de las mismas que nos han llegado pertenecen casi por entero a la arqueología. Por haberse encontrado muy pocas ruedas se infirió que existían pocas. Pero la revisión actual, de la documentación literaria y de la arqueológica, indica que las hidráulicas se generalizaron en el siglo II d.C.

La revisión se basa, por una parte, en el descubrimiento de más ruedas y, por otra, en el argumento no menos convincente de la analogía. De la Inglaterra del siglo XI, por ejemplo, apenas si queda rastro de una docena de molinos de cualquier clase, y, sin embargo, el catastro de Guillermo el Conquistador (llamado *Libro del Día del Juicio Final*) registra la existencia de por lo menos 5624. Si suponemos que la proporción entre los molinos existentes en realidad y los descubiertos es la misma para el mundo clásico, algunas de las explicaciones aceptadas sobre la escasez de ruedas hidráulicas empiezan a sonar a hueco.

Según la explicación recibida, la cuenca mediterránea, y en particular Italia, carecía de corrientes fluviales que ofreciesen el caudal constante en el que fían los molinos. Esta teoría sólo resulta convincente si uno supone que existían pocos molinos. Si uno admite, en cambio, que había muchos pronto viene a la mente que si hay algo en que los romanos destacaran era su preocupación por el abastecimiento de agua, ininterrumpido y seguro, mediante los acueductos. El acueducto debía ser ideal para las necesidades de un molino de agua, particularmente del modelo propulsado por arriba, como en el caso de Barbégal. Se puede aducir que el gasto era excesivo; pero una vez más, no falta la respuesta adecuada. De todos los usuarios del agua, los molinos constituyen un caso único, puesto que ni consumen agua ni la degradan. En teoría, por tanto, habría sido posible instalar molinos en cualquier punto del acueducto de una ciudad sin merma alguna de su función normal.

De hecho, algo de este género parece que se hizo en la ladera del Janículo, en los suburbios de Roma. Cabe preguntarse si éste fue un caso aislado o si los restos de otros molinos análogos han pasado inadvertidos porque nadie esperaba encontrarlos. Desde luego no era raro que los acueductos formasen cascadas en algún punto donde el terreno les hacía perder altura rápidamente, lo que facilitaría la construcción de un molino. Naturalmente, no en todos los casos hubiera sido económico. Un buen emplazamiento natural para un molino podía quedar separado de los

6. MECANISMO DE TRANSMISION necesario para convertir la rotación vertical de una rueda hidráulica, propulsada por arriba o por abajo, en rotación horizontal. En el dibujo que la recrea se supone que el transmisor era mayor que el piñón, con el fin de acelerar la velocidad del molino.

principales centros urbanos por una distancia excesiva o por una topografía inaccesible.

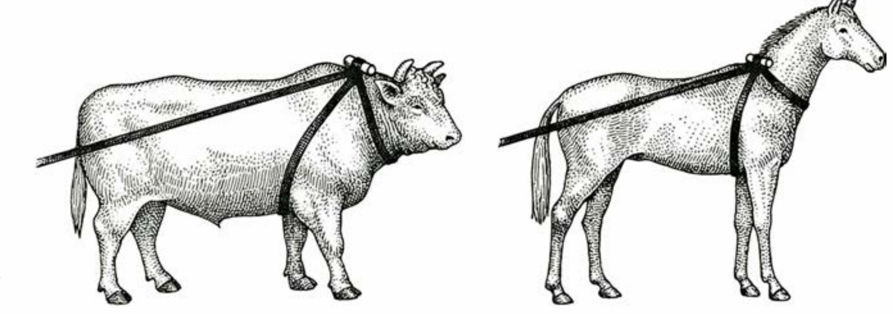
Pese a todo, no se puede negar que los romanos crearon menos molinos de lo que hubiese sido de esperar. ¿Por qué? La explicación tradicional –la teoría tecnológica de la decandencia romana– argumenta que carecieron de incentivos económicos para el desarrollo técnico mientras el Imperio estaba expandiéndose, y que ya habían perdido su capacidad técnica cuando las fronteras por fin se estabilizaron. Esta teoría se apoya casi por completo en la perniciosa influencia de la esclavitud.

Es peligroso echarle todas las culpas a la esclavitud. Cierto que el trabajo barato, sea libre o esclavo, puede tener un efecto estupefaciente en el desarrollo de técnicas tendentes a economizar mano de obra. Pero los esclavos solían ser más caros de lo que se cree, pues combinaban altas inversiones de capital con una vida activa corta (cuando se les abrumaba de trabajo) y con el alto coste de manutención en su período activo.

Los esclavos podían resultar económicos cuando irrumpían en el mercado en grandes cantidades, haciendo que bajara el precio medio de compra. Esta abundancia de esclavos se mantuvo mientras duró la expansión del Imperio y sus victoriosas legiones iban enviando a sus lares prisioneros de guerra desde las fronteras. Una vez que éstas se estabilizaron, en el siglo I d.C., y en mayor medida desde que los bárbaros comenzaron a ganar batallas, la oferta de esclavos se agotó y el precio de la mano de obra subió. Roma se encontró así en la situación de algunas empresas modernas: tenía que crecer o extinguirse.

Pero hay una razón mucho más directa por la que la rueda hidráulica languideció en la Antigüedad, y con ella construcciones como el molino de Barbégal y las culturas que podían depender de ellos. Es algo que tiene que ver con los atelajes de los caballos.

Aunque el empleo del caballo como animal de tiro estaba muy restringido por la carencia de herraduras, esa deficiencia era poca cosa en comparación con la inoperancia de las guarniciones en uso. Las correas de los arneses y la anatomía equina coincidían de tal manera, que si el caballo trataba de arrastrar una carga pesada lo único que conseguía era estrangularse. Para el transporte terrestre de mercancías –incluso por las excelentes vías romanas– los caballos estaban excluidos. La fuerza motriz la proporcionaban los bueyes, muy fuertes pero desesperadamente lentos.



7. EL ATALAJE ROMANO estaba pensado de modo que una yunta de bueyes pudiese tirar del yugo con sus cervices (*izquierda*). No servía para los caballos, que se estrangularían si trataban de tirar de una carga pesada (*derecha*). Los antiguos tenían que elegir, por tanto, entre bueyes penosamente lentos y caballos que desfallecían a causa del dolor. En cualquier caso, y de acuerdo con la hipótesis del autor, el transporte resultaba demasiado caro para justificar la producción en masa.

El transporte por tierra era, por tanto, de un precio prohibitivo para distancias superiores a los 25 o 35 kilómetros, que un buey tardaba varios días en cubrir. Esta limitación a la distribución de mercancías tenía una grave consecuencia: descartaba casi por principio las fábricas, pues el ahorro que comporta la producción en masa quedaría más que contrarrestado por el gasto de transporte de las mercaderías. Y sin producción en masa, poco incentivo cabía esperar para la mecanización.

Por consiguiente, la maquinaria impulsada por agua no era rentable, a menos que el agua fuera fácil de obtener a bajo precio, y que el mercado local fuese el adecuado. Otros mercados más distantes sólo podían ser abastecidos en los contados casos en que los molinos se pudieran construir cerca de algún puerto donde atracaran barcazas fluviales o barcos de transporte marítimo –con mucho, los medios de transporte más baratos de la Antigüedad– que cargaran la producción y la llevaran a su destino. El vate Ausonio, del siglo IV, describe una empresa de tal género –un aserradero de mármoles en el río Ruwer, afluente del Mosela, cerca de Tréveris. Es uno de los poquísimos molinos de los que se sabe que producían algo que no fuera harina.

¿Es acaso Barbégal único? No del todo. Se conocen otros dos molinos múltiples, pero ninguno de proporciones equivalentes. Uno está en Chemtou, en el oeste de Túnez, donde una combinación de puente y presa que data de comienzos del siglo II d.C. atravesaba el río Medjerda. Encajadas en un estribo del puente había tres ruedas de molino horizontales, una al lado de otra, en paralelo y no en serie como en Barbégal. Las ruedas estaban alojadas en recintos de piedra cilíndricos, conocidos como paraderas; la corriente del agua se recogía allí un poco descentrada, de modo que actuase principalmente sobre un lado de la rueda, como en una turbina. El otro molino está en Israel, en una presa del río Cocodrilo, cerca de la antigua Cesarea, a mitad de camino entre Haifa y Tel Aviv. Aquí también había dos ruedas horizontales, cada una de ellas en el fondo de una paradera, de modo que el agua, que entraba por arriba, golpeaba la rueda con su impulso incrementado por la caída. Ninguna de estas instalaciones ha sido estudiada a fondo, pero siguen siendo los únicos parangones conocidos de Barbégal.

Y queda la ironía. Barbégal está situado en una zona densamente poblada de un país moderno, desarrollado. No estaba enterrado ni requirió excavación más allá de una ligera limpieza. No era fácil que el molino pasase inadvertido; pues bien, a pesar de ello, permaneció en el olvido durante siglos.

La lección es clara. Otros Barbégal nos están esperando en las partes más remotas y menos estudiadas del Imperio. Si éste pudo pasar inadvertido hasta 1940 en Francia, ¿qué obras maestras no pueden permanecer aún ocultas en Irak o el norte de Africa, donde las arenas del desierto arropan hoy los restos de ciudades romanas? Merece la pena buscarlos para bien de la historia de la tecnología y la ciencia, la industria, la economía y la cultura.

BIBLIOGRAFIA COMPLEMENTARIA

THE ORIGINS OF WATER POWER: A PROBLEM OF EVIDENCE AND EXPECTATIONS. N.A.F. Smith en *Transactions of the Newcomen Society,* vol. 55, págs. 67-84; 1983-84.

SIFONES EN LOS ACUEDUCTOS ROMANOS en *Investigación y Ciencia,* n.º 107; págs. 80-86; agosto de 1985.

ROMAN AQUEDUCTS AND WATER SUPPLY. A. Trevor Hodge, Gerald Duckworth & Co., Ltd. (en prensa).

Orquídeas unisexuales

Dotadas de amplia variedad de formas, tamaños, colores y fragancias, las orquídeas poseen un refinado comportamiento reproductor que el autor pone aquí de manifiesto a través del género Catasetum

Gustavo A. Romero

"Si nos pidieran seleccionar la más interesante entre la multitud de tribus vegetales, deberíamos... quizás, estar dispuestos a darle preferencia al orden natural de las orquídeas. Si consideramos la elegancia general de los individuos, la durabilidad de las flores, los espléndidos colores, los deliciosos perfumes, o su extraordinaria estructura, sería difícil seleccionar cualquier otro orden superior a las orquídeas... y pocos aún igual a ellas."

John Lindley. (*Collectanea Botanica*, lámina 2, 1820).

Las orquídeas, quizá las flores más populares y mejor conocidas del mundo, han sido objeto de admiración de recolectores, cultivadores y científicos. Pero, ¿qué es una orquídea? Los miembros de esta familia se distinguen por compartir varios caracteres; a saber: el pistilo y los estambres están, al menos parcialmente, soldados y forman la llamada columna, un órgano compuesto; presentan, por lo general, una sola antera fértil, que contiene un gran número de granos de polen, conglomerados en masas compactas en la mayoría de los casos; el ovario aloja, a su vez, múltiples óvulos que, de ser polinizados, dan origen a semillas muy pequeñas; por último, uno de los pétalos, el labelo, ha adquirido un alto grado de diferenciación.

GUSTAVO A. ROMERO ha venido desempeñando el cargo de investigador asociado al Herbario de la Universidad de Harvard, en el que es responsable del herbario de orquídeas de Oakes Ames, desde 1988. Fue jefe fundador de la Estación Experimental Amazonas del Fondo Nacional de Investigaciones de Venezuela, entre 1981 y 1984, y director en 1987. Se doctoró en ecología y biología evolutiva por la Universidad de Indiana. Es asesor del gobierno de su país, Venezuela, en la creación de un Centro de Investigaciones Ambientales en la Amazonía.

Aunque las orquídeas crecen en todos los ecosistemas terrestres con la excepción de los polos, tienen su mayor representación en los trópicos. Se pueden encontrar sobre la tierra (terrestres) o, en contados casos, en el subsuelo (subterráneas), sobre los árboles (epífitas) o sobre piedras (litófitas). El hábito terrestre es la forma de vida más primitiva de la familia y la que predomina en las especies de clima templado; el epífito se considera la forma de vida más avanzada y la que predomina en las orquídeas tropicales.

La familia monopoliza los superlativos en la naturaleza: presenta una amplia variedad de formas, tamaños (desde plantas microscópicas en el género *Platystele* hasta plantas trepadoras de más de 10 metros de longitud en el género *Vanilla*), colores de la flor (del blanco al morado obscuro), fragancias (desde deliciosas hasta nauseabundas) y sistemas de polinización (desde moscas microscópicas hasta aves). Creíase, hasta hace poco, que constituía la familia de plantas con el mayor número de géneros y especies, con cerca de 700-800 géneros y 20.000-22.000 especies. Si bien los cómputos recientes indican que la primacía en abundancia corresponde a la familia de las Compuestas o Asteráceas, la del girasol, que tiene alrededor de 3000 géneros y 30.000 especies, pocos se atreverían a cuestionar la preeminencia de las orquídeas en belleza y complejidad de sus flores y sistemas de polinización.

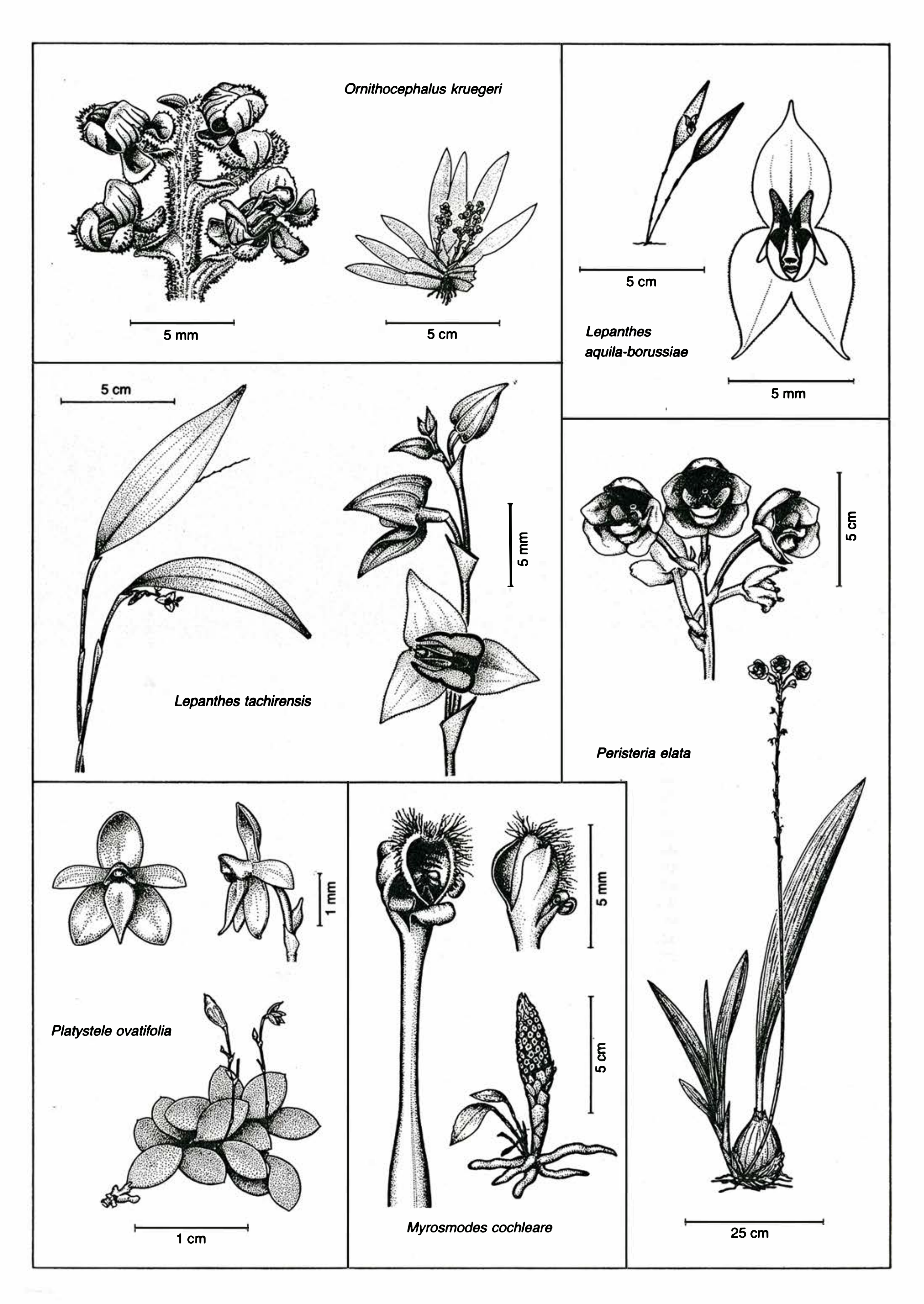
En efecto, quizás uno de los rasgos más distintivos de las orquídeas sean sus sistemas de polinización. Presentan polinización zoófila, no anemófila; es decir, deben ser necesariamente polinizadas por animales, ya que la naturaleza de su polen impide el transporte de éste por el viento. Según la especie que consideremos, las flores atraerán moscas, mosquitos, avispas, abejas, mariposas, polillas o aves. A diferencia de lo que ocurre en la mayoría de las plantas, en las orquídeas el polen nunca se ofrece como recompensa a los polinizadores; más aún: en un gran número de especies, ni siquiera se aporta ningún tipo de recompensa. Las orquídeas que sí la ofrecen producen néctar, aceites o, como veremos más adelante, fragancias. Entre las especies que no presentan ninguna contraprestación encontramos flores que se aprovechan de sus polinizadores por medio de algún mimetismo. Este señuelo se evidencia a través de dos formas: imitando a otra flor más común que sí ofrece néctar u otro tipo de recompensa o imitando a la hembra de la especie del insecto polinizador. En el segundo caso, la mímesis adquiere a veces un refinamiento tal, que la combinación de estímulos olfativos y táctiles induce a la pseudocopulación.

Las variaciones en tamaño, color, olor y forma entre especies de orquídeas son importantes indicadores del polinizador que utilizan. Así, las flores polinizadas por moscas y mosquitos suelen ser pequeñas y pueden producir, en la mayoría de los casos durante el día, desde olores muy dulces hasta nauseabundos que recuerdan el olor a la carne en descomposición. Las polinizadas por polillas en general son blancas y producen fragancias después del anochecer. Las polinizadas por mariposas y picaflores presentan a menudo flores rojas o anaranjadas con una guía muy conspicua en el labelo para la probóscide (el alargado aparato bucal de las mariposas y polillas) o el pico del polinizador.

A pesar de que las flores de las orquídeas demuestran una variación muy amplia, sus partes derivan de órganos florales básicos: sépalos, pétalos, estambres y carpelos. Además, aunque atraen una gama muy amplia de polinizadores, la familia es bastante uniforme en lo concerniente a su sexualidad. Más de un 95 por ciento de las orquídeas son hermafroditas (estambres y pistilos en la misma flor)



1. LA TRIBU CATASETINAE está constituida por el género *Catasetum* y otros afines, representados aquí en algunas de sus especies: *Clowesia warscewiczii*, *Cynoches haagii*, *Dressleria, sp.* y *Mormodes aromatica*, de arriba abajo en el sentido de las agujas del reloj. Defínese dicha agrupación de géneros por la presencia de un mecanismo activo para la colocación del polen en el polinizador y la producción de flores protrándricas o unisexuales.


Ornithocephalus kruegeri
5 mm
5 cm
5 cm
Lepanthes
aquila-borussiae
5 mm
5 cm
5 mm
Lepanthes tachirensis
5 cm
Peristeria elata
1 mm
Platystele ovatifolia
1 cm
5 mm
5 cm
Myrosmodes cochleare
25 cm

y homógamas (estigma y polen maduran simultáneamente). De ese modo, en una sola visita un polinizador puede dejar su carga de polen en el estigma y tomar el polen de la misma flor.

Otro grupo de orquídeas, mucho más reducido en número de géneros y especies, produce flores en las que el polen madura antes que el estigma; en estas flores protándricas, así se califican, el estigma puede recibir polen sólo después de habérsele extraído a la masculina su propia carga de polen. Un último grupo, más reducido, ha llegado al extremo de producir flores estrictamente masculinas o femeninas; es decir, que sólo producen, o sólo aceptan, polen. (En sentido riguroso, únicamente la generación de células haploides —los óvulos y los granos de polen o gametofitos— es femenina o masculina, pero no la planta —el esporofito— ni la flor; aquí, sin embargo, por razones de simplificación, diremos que una flor se considera masculina cuando sólo produce polen y, femenina, si sólo produce óvulos, semillas y frutos.) El género *Catasetum* Rich. se encuentra en este postrer grupo. Las flores unisexuales de *Catasetum* presentan una oportunidad única para estudiar la evolución de ciertas estrategias reproductivas en este grupo de orquídeas masculinas y femeninas.

El género *Catasetum* fue descrito por el botánico francés Louis Claude Richard y publicado *post-mortem*, en 1822, en la obra *Synopsis Plantarum* de Humboldt, Bonpland y Kunth. El nombre procede de las raíces griegas *katá* (abajo) y *seta* (cerda), en alusión a las antenas de las flores masculinas. En el momento de publicarse el género sólo se describieron dos especies: *C. maculatum* Kunth y *C. macrocarpum* Rich. ex Kunth. El siguiente botánico en interesarse por el género fue John Lindley, taxónomo inglés que, entre 1824 y 1847, publicó más de 30 especies.

Entre el extenso material de orquídeas que Lindley examinó entre 1830 y 1840 se encontraban dos especímenes llegados de Brasil que, aunque muy cercanos a *Catasetum*, para él diferían lo suficiente como para erigir dos nuevos géneros: *Myanthus* y *Monachantus*. Hubo que esperar, sin embargo, a finales del siglo pasado, y sólo después de las cuidadosas observaciones del propio Lindley, de Robert Schomburgk, de Charles Darwin y de Robert A. Rolfe, para poder establecer que Lindley había creado dos géneros superfluos basados en el dimorfismo sexual de *Catasetum*. Darwin, siguiendo sus propias observaciones y las de Lindley y Schomburgk, concluyó que *Monachanthus* era simplemente una flor femenina de *Catasetum* y que *Myanthus* era una flor hermafrodita. Rolfe, después de examinar los datos aportados por Darwin, determinó con acierto que este último había cometido un error en cuanto a *Myanthus*; en opinión de Rolfe, *Myanthus sensu* Darwin representaba otra flor masculina de *Catasetum* y, tal como había propuesto Darwin, *Monachanthus* representaba la flor femenina. (Lo veremos después, las flores masculinas varían mucho entre especies en color y morfología, mientras que las femeninas son bastante uniformes.) Esta hipótesis propuesta por Rolfe en 1891 es ampliamente aceptada hoy en día.

El género *Catasetum* es estrictamente neotropical, encontrándose desde México hasta el norte de Argentina y el sureste de Brasil. Hasta ahora se conocen entre 60 y 80 especies, con la mayor diversidad en Brasil (50-60 especies) y en Venezuela (25-30 especies). En su distinta presentación de plantas terrestres, litófitas o epífitas, coloniza una variedad amplia de ambientes en un intervalo altitudinal que comprende desde manglares al nivel del mar hasta las montañas de arenisca de la Guayana, a más de 2500 metros, aunque la mayoría de las especies no superan la cota de los 1000 metros.

Los géneros más afines a *Catasetum* son *Clowesia* Lindley, *Dressleria* Dodson, *Mormodes* Lindley y *Cycnoches* Lindley, que, juntos, constituyen la tribu Catasetinae. Esta subtribu se define por la presencia de un mecanismo activo para la colocación del polen en el polinizador y la producción de flores protándricas o unisexuales. (Fuera de esta subtribu, en las orquídeas sólo se conocen flores unisexuales en una especie endémica de China del género *Satyrium*.) La importancia de estos caracteres resalta al estudiar las interacciones entre plantas y sus polinizadores y la evolución de estrategias reproductivas. Sin embargo, para apreciar la complejidad y singularidad de la biología reproductiva de *Catasetum* primero hay que familiarizarse con los componentes del sistema: su morfología floral, amén de sus interacciones con el ambiente y sus polinizadores.

La flor del antecesor hipotético de las orquídeas estaba constituida por el pedicelo, que es su soporte, y por cinco verticilos, que son grupos de tres o más estructuras del mismo tipo dispuestos concéntricamente y que surgen al mismo nivel sobre un eje. Los cinco verticilos, comenzando por la parte basal de la flor, son: los sépalos, los pétalos, dos juegos de tres estambres cada uno (los órganos sexuales masculinos) y tres carpelos (los órganos sexuales femeninos). Los principales cambios que se observan en las orquídeas actuales, con respecto a su antecesor hipotético, han ocurrido en uno de los pétalos y en los órganos sexuales masculinos y femeninos.

Los tres sépalos de las orquídeas protegen a la flor durante su desarrollo. De colores llamativos en general, alcanzan un tamaño y una forma similares, aunque por su posición podemos distinguir dos sépalos laterales y uno dorsal. Los pétalos tienen también colores vistosos; destaca, entre ellos, el labelo o pétalo medio (opuesto al sépalo dorsal), que desempeña un papel importante en la atracción de los polinizadores y cuya morfología y tamaño varía de una especie a otra: desde semejar una plataforma de aterrizaje o un elaborado tobogán hasta llevar diversos ornamentos o remedar la figura de un insecto que atraiga a otros para inducir a la copulación. El labelo puede tomar dos orientaciones: hacia abajo, cuando la flor es resupinada, o hacia arriba, cuando es prona. El labelo de *Catasetum* varía de acuerdo con la sexualidad de las flores. La mayor diferencia en tamaño, forma, color y orientación se produce entre las flores masculinas, y la menor entre las femeninas.

Los dos verticilos siguientes corresponden a los estambres. Cada estambre está constituido, a su vez, por un filamento y una antera; la antera consta de polen y cubierta. Del antepasado hipotético a las orquídeas actuales se ha evolucionado hacia la reducción del número de estambres. Hay sólo dos géneros que mantienen los dos juegos de estambres (aunque reducidos en número total de los seis estambres originales a tres o dos); otros cuatro géneros perdieron un juego de estambres y sólo mantienen dos del verticilo interior. La mayoría

2. VARIABILIDAD FLORAL de las orquídeas. Más de un 90 por ciento de las orquídeas son hermafroditas y homógamas, es decir, flores en las que el estigma madura al mismo tiempo que el polen; ello permite que el polinizador deje su carga polínica en el estigma y tome el polen de la misma flor en una sola visita. Otro grupo de orquídeas, mucho más reducido en número de géneros y especies, produce flores en las que el polen madura antes que el estigma; nos referimos a las orquídeas protándricas. Por último, un tercer grupo, aun más restringido, produce flores estrictamente masculinas o femeninas, es decir, sólo producen o sólo aceptan polen; a este grupo pertenece el género *Catasetum*.



3. TODAS LAS ORQUIDEAS comparten varios caracteres. Estambres y pistilo están, al menos parcialmente soldados y forman la columna, que es un órgano compuesto. Las orquídeas presentan por lo común una sola antera fértil, que contiene un gran número de granos de polen, conglomerados en masas compactas en la mayoría de los casos. El ovario aloja múltiples óvulos que, tras su fecundación, originan semillas muy pequeñas. Uno de los pétalos, el labelo, ha adquirido un alto grado de diferenciación. Las flores unisexuales de *Catasetum* ofrecen una oportunidad única para estudiar las estrategias reproductivas de las orquídeas. Las fotografías pertenecen a flores masculinas de *C. bicolor*, izquierda, y *C. cristatum*, derecha.

de las orquídeas, sin embargo, sólo mantienen un estambre del verticilo exterior.

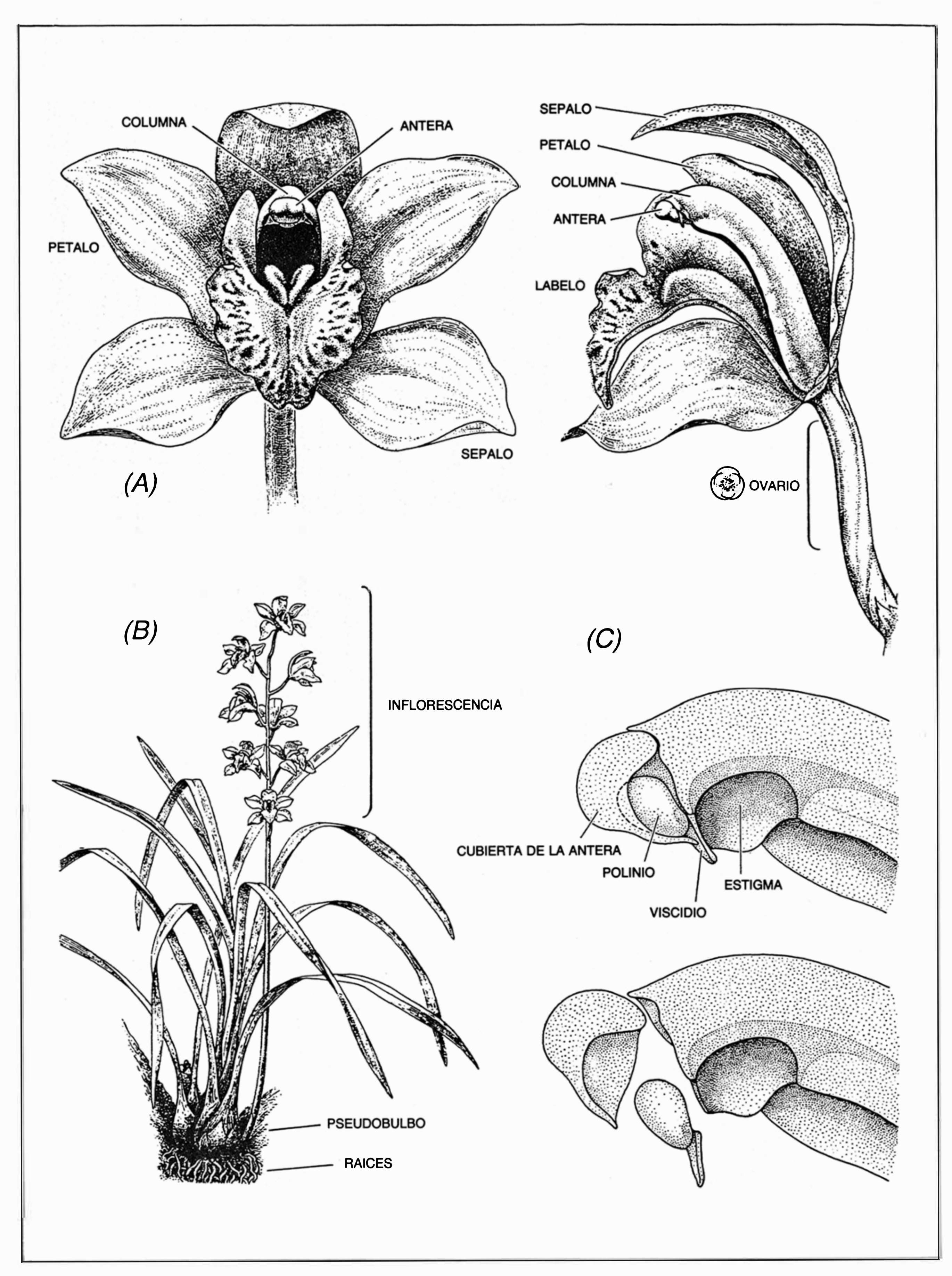
El polen ha experimentado una serie de modificaciones en el curso de la evolución y, aunque sea en raras ocasiones un polvo pegajoso, habitual en otras familias, se conglomera por lo común en masas granulosas o compactas, céraceas, de un número de unidades discreto que varía entre dos y ocho, en cuyo caso se les denomina polinios. Los polinios pueden estar asociados con otros órganos, tales como las caudículas, el estípite y el viscidio. El conjunto formado por polinio, caudícula, estípite y viscidio se llama polinario. Las caudículas son estructuras elásticas, derivadas de la antera, que conectan los polinios con el estípite o directamente con el viscidio si no hay estípite. Este conecta las caudículas con el viscidio, que posee un pegamento de secado rápido que adhiere el polinario al polinizador. Las caudículas proceden de la antera; el estípite y el viscidio, del estigma. Fuera de las orquídeas, tan sólo las Asclepiadáceas presentan también polinios.

Y llegamos al verticilo más interno, el pistilo. Este consta de tres carpelos y se divide en ovario, estilo y estigma. Los tres carpelos se hallan soldados en una sola unidad. El ovario posee, a su vez, un solo compartimiento o lóculo, aunque en unos pocos géneros (verbigracia, *Apostasia, Neuwiedia, Phragmipedium* y *Selenepedium*) se aprecian tres compartimientos. Además, el ovario se encuentra fusionado con la parte basal de los verticilos exteriores (sépalos, pétalos y estambres) y toma una posición inferior con respecto a éstos.

El estilo y el estigma, por su parte, pueden fusionarse con los estambres y formar la columna o ginostermo; fusión que puede ser parcial o total. Es parcial en los géneros *Phragmipedium* y *Apostasia*, donde se distinguen dos o tres anteras con vestigios de los filamentos y el estigma con vestigios del estilo. Es total en el resto de las orquídeas, y entonces resulta imposible distinguir, en la flor madura, entre componentes originales y columna.

En el género *Catasetum*, la columna varía con el sexo de las flores. En las masculinas, la columna remata en una antera que porta los polinios bien desarrollados, igual que las estructuras asociadas a los mismos, es decir, las caudículas, el estípite y el viscidio. Comparado con el de otras orquídeas, aquí el viscidio se distingue netamente. El estípite, por su parte, se ve sometido a tensión durante el desarrollo de la columna, esfuerzo que se hace máximo en el momento de abrirse la flor.

La columna de las flores masculinas *Catasetum* tiene, hacia la base, un par de antenas derivadas de la torsión de las paredes laterales del estigma, que funcionan a modo de gatillos o dis-



4. CARACTERES PRINCIPALES de una orquídea. La aquí ilustrada pertenece al género *Cymbidium*. En a, se ofrece la visión frontal y lateral de un corte. En b, se representa la planta entera en un ambiente epífito. Especies de este género medran también en el suelo. Por último, en c, se recrean algunos pormenores de la columna, estructura que sirve para distinguir entre las orquídeas y otras plantas asimismo monocotiledóneas.

paradores para la expulsión del polinario. La estimulación de una de las antenas provoca la descarga violenta e instantánea de la tensión del estípite con la expulsión consiguiente del polinario y de la cubierta de la antera. En las flores femeninas, la columna es mucho más corta y engrosada que en las flores masculinas; la antera y el polinario se desarrollan sólo parcialmente y se caen al poco tiempo de abrirse la flor. En contraste con las masculinas, las flores femeninas de *Catasetum* carecen de antenas y muestran una cavidad estigmática funcional; la apertura de esta cavidad es una ranura semilunar.

Los frutos de la mayoría de las orquídeas son cápsulas, frutos simples, secos que se abren espontáneamente al llegar a la maduración, si exceptuamos algunas especies de frutos carnosos y semillas asociadas con una pulpa. (Entre ellas, las especies del género *Vanilla*, de las que se extrae la vainilla.) El número de semillas por cápsula varía, a tenor de la especie, entre unos pocos miles y más de un millón, abundancia que caracteriza al género *Catasetum* y sus afines. Las semillas acostumbran ser muy pequeñas (de 0,3 a 0,5 milímetros) y livianas (de 0,3 a 14 microgramos); las dispersa el viento.

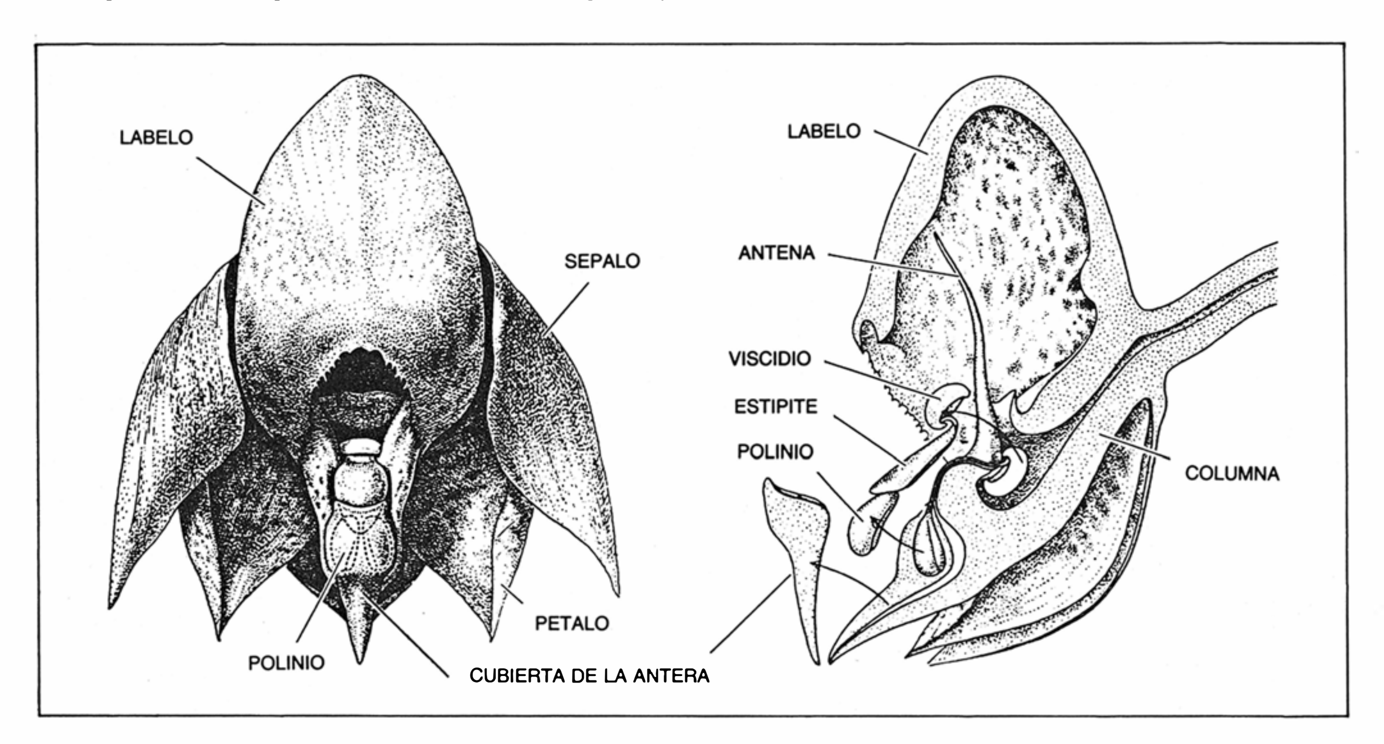
En las plantas de *Catasetum* se desarrolla, cada año, un pseudobulbo con sus hojas y su sistema radicular. La temporada de crecimiento suele corresponder con el período del año de mayor precipitación. Al final de la temporada de lluvias, las hojas caen y muere el sistema radicular; las reservas de agua, energía y nutrientes que ha almacenado el pseudobulbo mantienen a la planta durante la época de sequía. Las inflorescencias, que pueden ser de una a cinco, surgen en cualquier momento de la temporada de crecimiento, según las especies.

La polinización corre a cargo de abejas de la familia Apidae, en particular de los géneros *Eufriesea*, *Euglossa* y *Eulaema*. Las orquídeas que atraen a estas abejas no ofrecen néctar y sólo reciben la visita de los machos; las hembras son atraídas por otras plantas que ofrecen néctar o polen, o ambos, en recompensa. A los machos les atraen compuestos terpenoicos y aromáticos, en su mayoría volátiles y fragantes. Con los tarsos delanteros recolectan estas substancias y las almacenan en las tibias traseras. Se ha propuesto que, luego de sufrir ciertas modificaciones químicas, los compuestos son utilizados en la atracción de las hembras; pero no disponemos de datos suficientes que corroboren esa hipótesis. Creíase antaño que las abejas "se drogaban" o "emborrachaban" mientras recolectaban fragancias en orquídeas, estado del que se aprovechaban las orquídeas para dominar al polinizador. Opinión que hoy se ha visto carece de fundamento.

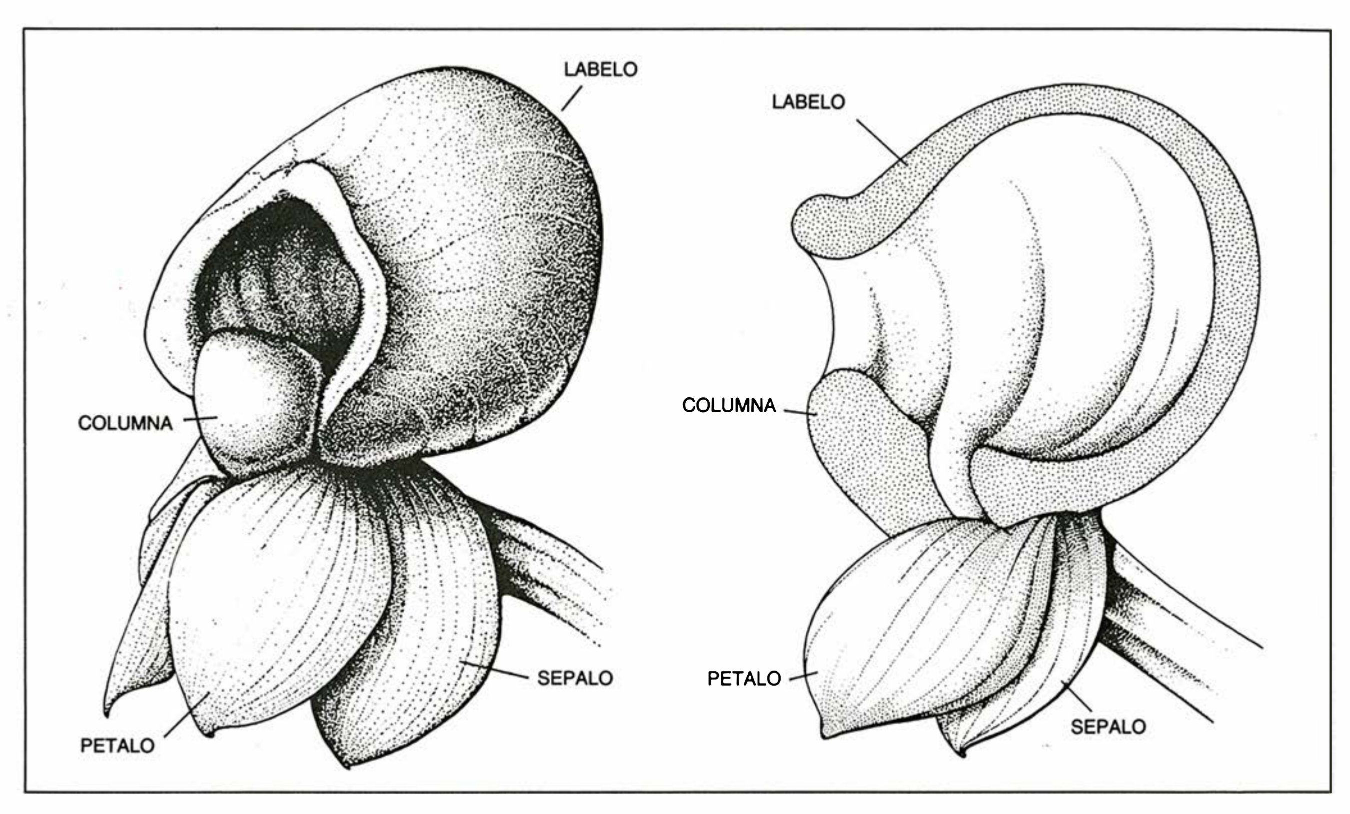
Las abejas que polinizan a *Catasetum* pertenecen a especies de los géneros *Euglossa* y *Eulaema*. Las flores masculinas y femeninas de estas orquídeas, aunque dotadas de un manifiesto dimorfismo sexual, producen la misma fragancia. El proceso de polinización se inicia cuando la abeja, atraída por esta fragancia, se aproxima a la flor masculina, inspecciona la flor y aterriza en el labelo. El artrópodo penetra, en posición normal, en el labelo si éste se halla abajo, o lo hace desde abajo (patas arriba) si el labelo está encima.

Las células que producen la fragancia, los osmóforos, residen en el interior del labelo. La abeja procede a recolectar fragancias en esa región, acercándose paulatinamente a las antenas. Tarde o temprano toca una de ellas, provocando así la expulsión instantánea del polinario y de la cubierta de la antera; este conjunto se coloca violentamente en su dorso, donde se adhiere con el pegamento viscoso del viscidio. Sacudida o asustada por esta acción, la abeja toma vuelo y se retira de la flor. La cubierta de la antera se reseca y cae a los pocos minutos, dejando al descubierto los dos polinios.

La abeja, atraída por la misma fragancia, se aproxima a la flor femenina y entra al revés en el labelo. (Las flores femeninas son siempre pronas, es decir, con el labelo arriba.) El polinario, que está abisagrado entre el viscidio y el estípite, oscila con libertad en el dorso de la abeja. Esta entra y sale de la flor conforme va recolectando fragancias y las transfiere a las tibias traseras mientras revolotea delante de la flor. En una de estas sali-



5. FLOR MASCULINA y mecanismo disparador del proceso de polinización en una orquídea del género *Catasetum*. La visión frontal de la izquierda expone la situación del polinio que subyace bajo la cubierta de la antera. El corte lateral de la derecha nos abre el interior de la flor. Las abejas, al tocar una de las antenas, liberan el viscidio pegajoso junto con el estípite y los polinios y se llevan este conjunto adherido al cuerpo.



6. FLOR FEMENINA del género *Catasetum*. Las flores femeninas de estas orquídeas poseen una columna más corta y engrosada que las flores masculinas, carecen de antenas y muestran una cavidad estigmática funcional; la apertura de esta cavidad es una ranura en forma semilunar.

das, uno de los polinios se aloja en el estigma y completa así el proceso de polinización.

Cuando la planta sólo produce flores masculinas, su gasto en energía reproductiva (inversión reproductiva) se limita a la producción y mantenimiento de flores y polen; en el caso de las flores femeninas esta inversión puede ser mucho mayor. Primero, las flores femeninas duran mucho mucho más que las masculinas (una o dos semanas frente a cuatro o cinco días) y, segundo, si una de ellas es polinizada, la planta tiene que desarrollar y mantener el fruto con todas sus estructuras asociadas hasta que maduren las semillas. Este período puede durar hasta un año y coincidir con la temporada de sequía.

En conjunto, el desarrollo y mantenimiento del fruto implica un gasto mucho más alto de agua, energía y nutrientes que lo invertido por las plantas masculinas. En particular, el gasto en nutrientes crece con el número de semillas por fruto y en virtud de otro factor: en las semillas, carentes de endospermo, una alta proporción de su peso lo constituyen los elementos nitrógeno y fósforo. Añádase a ello que el análisis de los nutrientes pone de manifiesto que, durante la maduración de los frutos, las plantas invierten en las semillas un alto porcentaje del total de su disponibilidad de estos dos elementos. La diferencia en los costos de producción entre polen y semillas puede ser tan desproporcionada que algunas plantas, aunque pueden producir polen, no podrían producir frutos por escasez de recursos (agua, energía o nutrientes); esta escasez puede estar relacionada con el tamaño de la planta o con la disponibilidad de recursos.

La diferencia entre gasto invertido en polen y el invertido en la producción de semillas ha constituido un factor determinante en la evolución de dos estrategias reproductoras de *Catasetum*: la unisexualidad y la facultad de poder cambiar de sexo entre estaciones. Estas estrategias son adaptativas en la medida en que permiten que una planta, con independencia de su tamaño y del sitio donde se desarrolle, puede canalizar sus recursos hacia el sexo que optimice su reproducción, y que, si llega a invertirlos en semillas en una estación, disminuyendo así sus reservas de agua, energía o nutrientes, pueda producir flores masculinas en la estación siguiente.

En *Catasetum*, además de la unisexualidad y del cambio de sexo, existe otro rasgo distintivo: el dimorfismo sexual, que puede llegar al extremo en que las flores no parezcan pertenecer a la misma especie y, ni siquiera, al mismo género. El dimorfismo se expresa en la forma y color del labelo y, de manera muy señalada, en la estructura de la columna. ¿Qué condiciones podrían haber propiciado la evolución de semejante dimorfismo sexual?

En *Catasetum* se pueden hacer dos observaciones muy importantes para elucidar la evolución del dimorfismo sexual. Primero, en la naturaleza encontramos muchas más flores masculinas que femeninas. Segundo, cada flor femenina puede recibir un solo polinio, ya que la ranura estigmática resulta demasiado estrecha para alojar dos y se cierra a las pocas horas de recibir el primer polinio. Ambas observaciones, sumadas al hecho de que cada flor masculina está capacitada para polinizar dos flores femeninas, abonan la existencia de una fuerte competencia entre las masculinas para acceder al limitado número de flores femeninas. Competencia que se agudiza si consideramos que, cuando una abeja que ha recibido un polinario visita otra flor masculina y recibe un segundo polinario, existe la probabilidad, cifrada entre un 33 y un 66 por ciento, de que el segundo polinario inutilice o bloquee mecánicamente el primero. Ante esas condiciones, cualquier rasgo que proporcione margen de ventaja a una flor masculina en su acceso a las femeninas se verá altamente privilegiado por la selección natural. De igual manera, cualquier rasgo que minimice el nú-

mero de polinarios de una inflorescencia que recalen en el mismo polinizador, evitando la competencia entre polinarios de una misma planta, se verá favorecido por la selección.

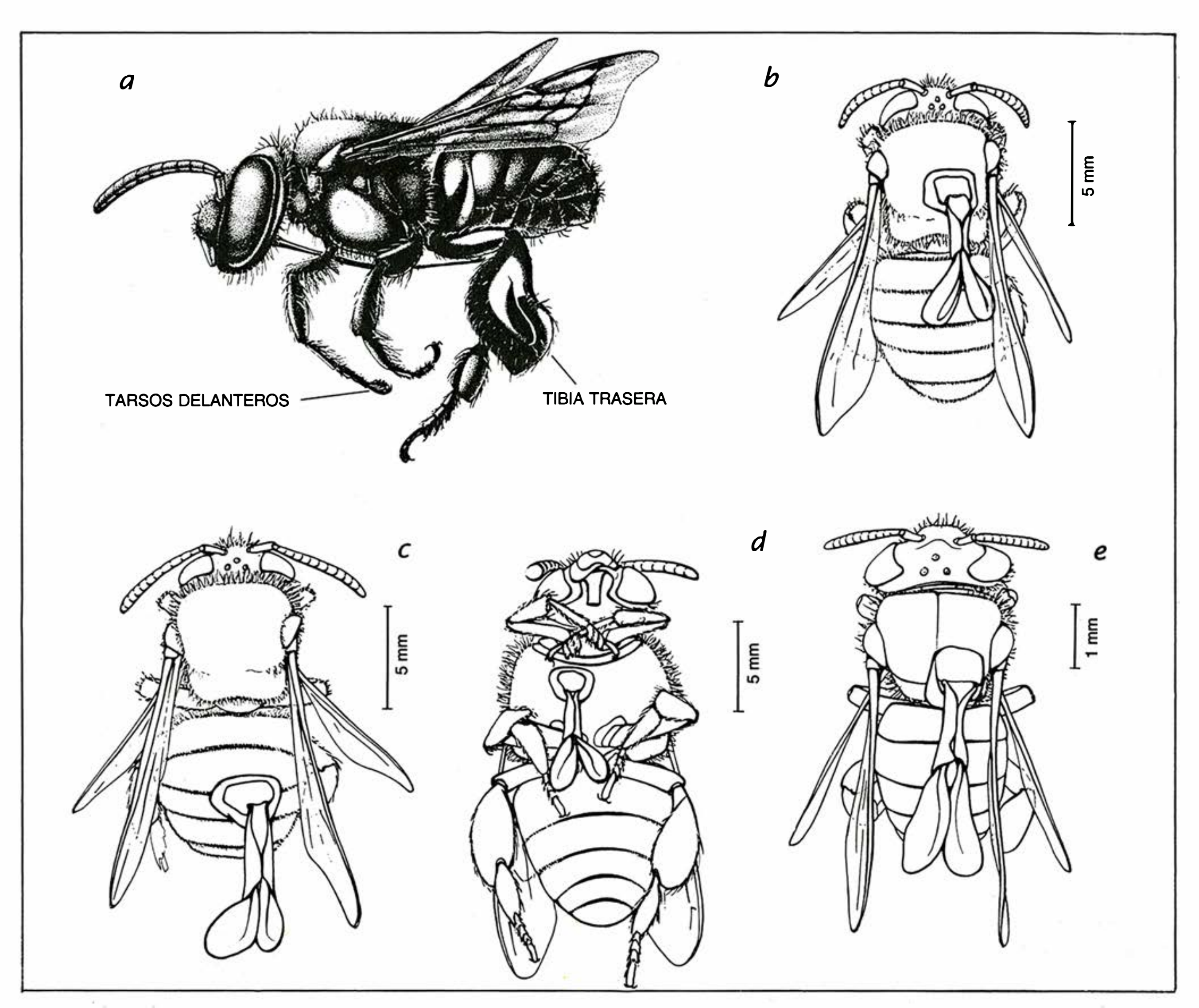
En las orquídeas del género *Catasetum*, la colocación violenta del polen en el animal polinizador parece haber seguido esta pauta evolutiva: las abejas que reciben el impacto del polen se alejan de la flor rápidamente y no regresan, que se sepa, a otra flor masculina. Dos tipos de pruebas prestan apoyo a la hipótesis. En primer lugar, la proporción de polinizadores que portan polinarios es mucho mayor entre los euglósidos que visitan las flores femeninas que entre los que acuden a las masculinas, a pesar de que se aproximan a unas y otras en proporciones similares. En segundo lugar, de las abejas con carga de polen que son atraídas hacia trampas de fragancias artificiales a lo largo de toda una temporada de floración, el número de abejas que tienen un solo polinio, o ninguno, es mucho mayor que el número de abejas con dos polinarios. De manera que las abejas que han recibido ya un polinario tienden a visitar más a las flores femeninas, lo que se manifiesta en la ausencia de uno o sus dos polinios, que a otras flores masculinas, que se evidenciaría con la presencia de dos polinarios. (Conviene tener bien clara la diferencia entre polinio, o polen agregado en masas compactas, y polinario, o polinios asociados con las caudículas, el estípite y el viscidio.) Sin embargo, la aversión de las abejas con cargas de polen a otras flores las mantendría alejadas de las flores femeninas, a menos que hubiera cierto dimorfismo sexual, rasgo que probablemente evolucionó en paralelo con el rechazo a un segundo polinario.

Entre las flores masculinas existe una gran diversidad en color y forma, incluso en el seno de la misma especie. Todo lo contrario que ocurre entre las flores femeninas: todas son verdes con el labelo en forma de casco. Esta homogeneidad morfológica de las segundas podría haber surgido, y persistido, como resultado de la complejidad del proceso de polinización en *Catasetum*. Cuando se observan polinizadores recolectando fragancias en las flores femeninas, se aprecia en seguida el distinto tiempo que tardan unas abejas y otras en hacerlo de la manera "correcta" para la flor, es decir, invertidas (o patas arriba), a fin de que el polinario se aloje en la ranura estigmática. Unas abejas atinan a la primera, mientras que otras tardan hasta 10 minutos en conseguirlo, señal de que existen diferentes niveles de experiencia entre esos polinizadores. Puesto que la abeja sobrevive a las estaciones de reproducción de varias especies, diríase que resulta beneficioso para la flor femenina aprovecharse del entrenamiento que las abejas obtienen en otras flores femeninas, promoviendo así la convergencia en la morfología de las flores femeninas.

Aunque la convergencia en la morfología de la flor femenina tal vez no promueva el aislamiento reproductor de especies potencialmente hibridizables (por su presencia en el mismo hábitat o por el tamaño de sus polinios), muchas especies logran ese aislamiento mediante la floración diferencial (unas florecen antes que otras), la atracción de polinizadores distintos o, si comparten polinizadores, la colocación de los polinarios en diferentes partes de la abeja.

Como se mencionó antes, diferentes especies de *Catasetum* pueden florecer en cualquier período entre el principio y el final de la temporada de crecimiento. Cuando concurren varias especies en una misma región, éstas tienden a florecer en temporadas diferentes. Otras especies han desarrollado, en su evolución, un espectro de fragancias o morfologías florales, o ambas, distintas del generado por otras con las que podrían hibridarse; por último, una especie puede colocar su polinario en el dorso del tórax, otra en el abdomen y otra en la parte ventral del tórax. A pesar de todas esas estrategias evolutivas que tien-

7. POLINIZACION de las orquídeas del género *Catasetum*. El proceso comienza cuando la abeja, atraída por la fragancia que despiden las flores, masculinas y femeninas, se aproxima a la flor masculina, inspecciona la flor y aterriza en el labelo. El artrópodo penetra, en posición normal, en el labelo si éste se halla abajo, o lo hace desde abajo (patas arriba) si el labelo está encima. Se acerca a las antenas y excita el polinario; el polen desprendido se adhiere a su abdomen. Luego, atraída por la misma fragancia, acude a una flor femenina para depositar el polen en el estigma.



8. POLINIZADORES de *Catasetum*. Las abejas que polinizan las plantas del género *Catasetum* pertenecen a especies de los géneros *Euglossa* y *Eulaema*. En *a* se muestra una vista lateral de *Euglossa gaianii* Dressler. En *b* se ofrece una vista dorsal de esta misma especie, mostrando un polinario de *Catasetum bicolor* Klotzsch. En *c*, *d* y *e* se muestran vistas de *Eulaema cingulata* mostrando polinarios de *C. pileatum*, *longifolium* y *macrocarpum*.

den a promover el aislamiento reproductor, hasta ahora se han detectado casos de hibridación natural entre una multitud de especies en Ecuador, Colombia, Brasil y Venezuela.

La descripción de la ecología y biología evolutiva del género *Catasetum* nos ha mostrado cómo, en ciertas orquídeas, diferentes presiones de selección sobre las funciones masculinas (la producción de polen) y femeninas (la producción de óvulos, frutos y semillas) pueden favorecer la evolución de estrategias reproductoras muy complejas.

Cabría, sin embargo y a modo de corolario, preguntarse por qué –si el contraste entre el costo de polen y semillas existe en muchas otras plantas y orquídeas– no presentan flores unisexuales y cambio de sexo entre estaciones todas las plantas y orquídeas.

La respuesta a esa cuestión ilustra ciertas características fundamentales de los procesos evolutivos en la naturaleza. La evolución de determinados caracteres puede ser aleatoria y hasta cierto punto errática; porque hay en juego muchas variables, las mismas presiones selectivas sobre diferentes plantas en el mismo medio pueden inducir la evolución de estrategias distintas. Además, estas diferentes estrategias pueden ser igualmente adaptativas, aunque tomen sus propios caminos para hacer frente a idénticas presiones selectivas. Considérese, por último, que muchas estrategias, caracteres y adaptaciones, como las que se vieron antes, pueden surgir por deriva genética, es decir, por la ocurrencia de cambios aleatorios en la frecuencia de genes en pequeñas poblaciones aisladas que no pueden atribuirse a la selección natural, a mutaciones o a migraciones. Cada estudioso de la evolución debe siempre mantenerla como explicación alternativa a sus propias hipótesis. Sólo así, a través de la comprensión del conjunto de procesos evolutivos, tarde o temprano podemos llegar a entender cabalmente de qué modo opera la evolución en la naturaleza.

BIBLIOGRAFIA COMPLEMENTARIA

POLLINATION AND VARIATION IN THE SUBTRIBE CATASETINAE (ORCHIDACEAE). C. H. Dodson en *Annals of the Missouri Botanical Garden*, vol. 49, págs. 35-56; 1962.

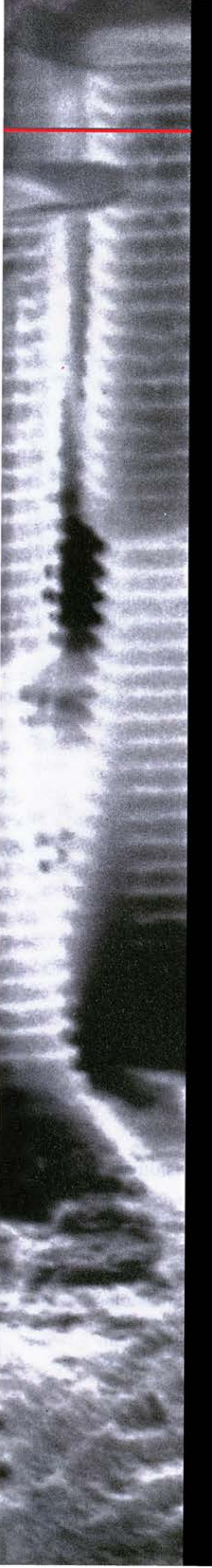
THE ORCHIDS. R. L. Dressler. Harvard University Press; Cambridge, 1981.

ORCHID FLORAL FRAGANCES AND MALE EUGLOSSINE BESS: METHODS AND ADVANCES IN THE LAST SESQUIDECADE. N. H. Williams y W. M. Whitten en *Biological Journal* (Lankaster), vol. 164, págs. 355-395; 1983.

SEXUAL DIMORPHISM IN *CATASETUM* ORCHIDS: FORCIBLE POLLEN EMPLACEMENT AND MALE FLOWER COMPETITION. G. A. Romero y C. E. Nelson en *Science*, vol. 232, págs. 1538-1540; 1986.

NANOTECNICA

Elizabeth Corcoran



La nueva generación de técnicas electrónicas y ópticas puede surgir de pequeños fragmentos de materiales que confinan a los electrones en espacios con menos de tres dimensiones.

LOS LASERES DE EMISION SUPERFICIAL de tamaño muy pequeño, con unas dimensiones del orden de seis micrometros de altura por uno de diámetro en la base, están constituidos por más de 30 pares de espejos de arseniuro de aluminio y arseniuro de galio. La luz se emite y amplifica mediante un pozo cuántico situado en el centro de la pila. Los espejos atrapan la luz y producen reflexiones múltiples en el interior de la pila.

Para preparar estos láseres, se operan unos cortes adecuados en una lámina de material semiconductor mediante una técnica de grabado por haces iónicos asistida químicamente. Antes de que los dispositivos que muestra la figura pudieran comprobarse, hubo que someter las estructuras a otros procesos; entre ellos, la instalación de unos contactos metálicos delgados en la parte superior de cada láser. Estos microláseres, construidos en la Bell Communications Research, emiten miliwatts de luz infrarroja a través del sustrato transparente que constituye su base.

Son las últimas horas de la tarde de un viernes cualquiera en Red Bank, New Jersey. Los pasillos de la Bellcore (Bell Communications Research) han quedado desiertos. Afanada en su laboratorio, una joven estudia un conjunto de gráficos que emergen en la pantalla de su ordenador. Detrás de ella, un haz luminoso de un verde azulado procedente de un láser de argón atraviesa un juego de lentes colocadas sobre una mesa; desaparece a través de un orificio practicado en un cilindro de acero inoxidable, que se mantiene lleno con el aire líquido que le llega por un tubo humeante y recubierto de hielo. Allí, en el interior de una cámara fría, a una temperatura inferior a los cinco grados Kelvin, el haz de láser arriba a su destino final: una partícula microscópica situada en el centro de una lámina cuadrada y plana que, como mucho, medirá contados milímetros. Excitada por el haz de láser verde-azulado, la partícula emite destellos brillantes y rojos.

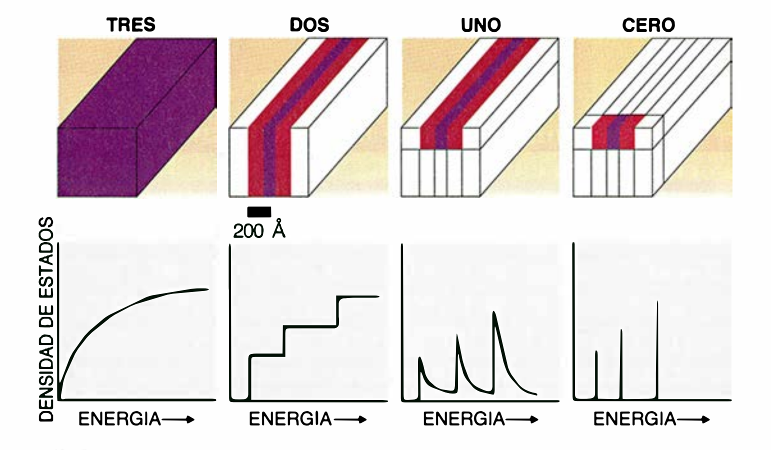
Este juego de luces interesa sobremanera a María Brasil, por más que la veamos de espaldas al mismo. ¿Cuánta luz emite la partícula? ¿Con qué longitud de onda? Las respuestas se esconden en la lenta acumulación de datos que aparecen en la pantalla del ordenador de Brasil. La partícula en cuestión emite luz con una frecuencia exactamente igual a la que predicen sus fabricantes.

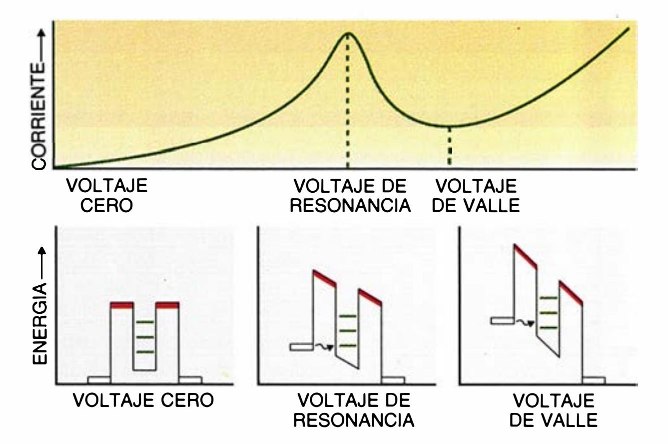
El pequeño emisor de luz roja producido por la compañía Bellcore, con un tamaño menor que el de un grano de arena, es precursor de una nueva generación de materiales electrónicos y ópticos que reciben la denominación común de dispositivos "cuánticos". A lo largo de estos últimos años, las técnicas de manipulación y observación de la materia a escala atómica han avanzado con velocidades asombrosas. En abril del año pasado, los científicos del Centro de Investigación Almaden de la empresa IBM, en San José, California, consiguieron deletrear el nombre de su compañía con átomos de xenón utilizando un microscopio de efecto túnel.

Aunque mover los átomos uno a uno constituye todavía un privilegio de laboratorio, la preparación de capas de extraordinaria finura se ha convertido ya en asunto muy serio. Controlando con precisión

REDUCIENDO LAS DIMENSIONES

La dimensionalidad de un material se puede reducir encerrándolo entre dos capas de otro material cuyos electrones posean mayor energía. Este confinamiento modifica la densidad de los estados electrónicos, o niveles energéticos específicos, que se llenarán con los electrones incidentes (*izquierda*). La corriente conducida por un dispositivo de pozo cuántico, representada por los niveles energéticos en verde (*derecha*), alcanza un valor máximo cuando el nivel de energía de los electrones incidentes coincide, o está en resonancia, con uno de los niveles energéticos del pozo cuántico. Para voltajes más altos o más bajos, la corriente que atraviesa el dispositivo es muy débil.





la estructura y la composición de capas de materiales cuyo espesor es sólo de uno o dos átomos, los científicos están demostrando que son capaces de programar las características electrónicas que desean en un determinado compuesto. "Es como tener en tus manos los botones de mando de la naturaleza", dice Mark A. Reed, profesor de ingeniería eléctrica de la Universidad de Yale.

La partícula roja de Bellcore, por ejemplo, es una estructura compleja constituida por capas de semiconductores de seleniuro de zinc y telururo de zinc. El espesor de cada capa es, a lo sumo, de 20 angstroms, un valor equivalente a algunas cienmilésimas del grosor de un cabello. Dado que dicho espesor es menor que la longitud de onda característica de los electrones, éstos quedan atrapados en dicha capa.

Cuando resultan confinados en tales capas, los electrones pierden un grado de libertad. Desde el punto de vista de los electrones, las delgadas capas del emisor luminoso de Bellcore se comportan como si sólo tuvieran dos dimensiones. Los científicos pueden reducir las dimensiones de los materiales todavía más. Realizando un corte muy delgado de este material, en una dirección perpendicular a los planos, se podría obtener un dispositivo unidimensional. Realizando un nuevo corte en este dispositivo, se llegaría a las últimas consecuencias en la reducción dimensional, obteniendo un punto con dimensión cero.

La reducción en el número de dimensiones de un material fuerza a los electrones a ocupar diferentes estados de energía. La creación de materiales en los que los electrones pongan de manifiesto dichos estados energéticos insólitos constituye el objetivo fundamental de la investigación de dispositivos cuánticos. Mediante el control del tamaño físico de esas estructuras, se provocan cambios previsibles en las energías de los electrones, para así obtener las propiedades electrónicas deseadas. Cuanto menor sea la dimensión, más fino resulta el ajuste. Un punto de dimensión cero, o punto cuántico, se puede considerar un átomo hecho a la medida.

Esta alquimia moderna abre el camino para crear toda una serie de dispositivos electrónicos y ópticos fundamentalmente diferentes. En vez de los conmutadores estrictamente binarios (con funciones de encendido/apagado) que los ordenadores utilizan, se confía en fabricar dispositivos conmutadores de función múltiple y montarlos para que operen en paralelo. El empleo de estos dispositivos permitiría fabricar modelos mucho más potentes de ordenadores lógicos y construir bloques de circuitos integrados muchísimo más pequeños y rápidos que los actuales. Algunos investigadores se hacen lenguas ante la posibilidad de construir un "superordenador sobre una sola pastilla".

Se mira también hacia la fabricación de láseres más finos y rentables, que permitirían enviar muchos más datos de lo que ahora es posible a través de redes de fibras ópticas. Gracias a ello, las empresas y los hogares contarían con una ingente cantidad de servicios de vídeo, ordenadores y telecomunicaciones. La feliz unión entre estos dispositivos electrónicos y los láseres convertiría en realidad un sueño acariciado durante mucho tiempo: la utilización de circuitos optoelectrónicos, en los que se integran electrones y fotones, permitiría fabricar componentes para ordenadores y redes de telecomunicaciones mucho más potentes y rápidos.

La empresa Bellcore no está sola en la investigación de estos dispositivos cuánticos. La estadounidense Fundación Nacional de la Ciencia ha creado un centro en la Universidad de California en Santa Bárbara dedicado en exclusiva al estudio del confinamiento cuántico. Las empresas IBM y AT&T Bell, entre otras, han aportado también su importante contribución. Los investigadores de Alemania, Inglaterra, Francia y, desde luego, Japón están también haciendo grandes avances en este campo. El tema clave con el que se enfrentan todos los investigadores consiste en averiguar de qué modo se pueden incorporar estos nuevos materiales —pozos cuánticos, alambres y puntos— en los dispositivos operacionales.

Jugando con las dimensiones

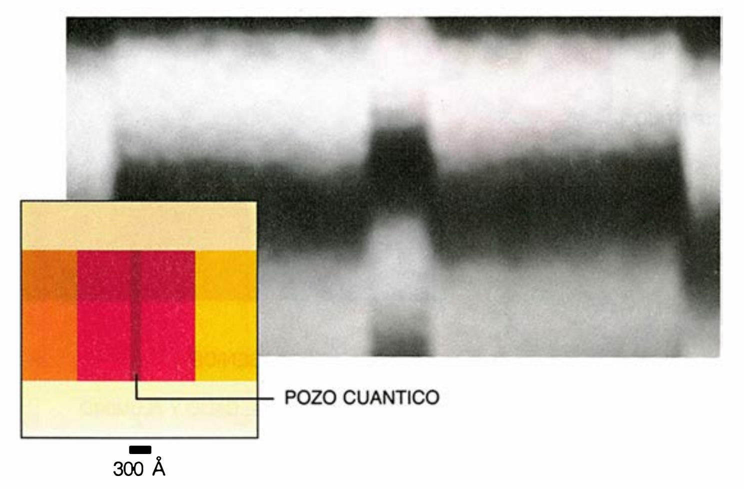
En el suelo del despacho de Leroy L. Chang, en el Centro de Investigación Thomas J. Watson de IBM, en Yorktown Heights, hay una estera tejida por varillas de bambú ensambladas en matriz: así se encontraban los átomos con los que él y sus colaboradores Leo Esaki y Raphael Tsu empezaron a jugar a finales de los años sesenta. Su objetivo era construir estructuras capaces de atrapar electrones en medios limitados dimensionalmente.

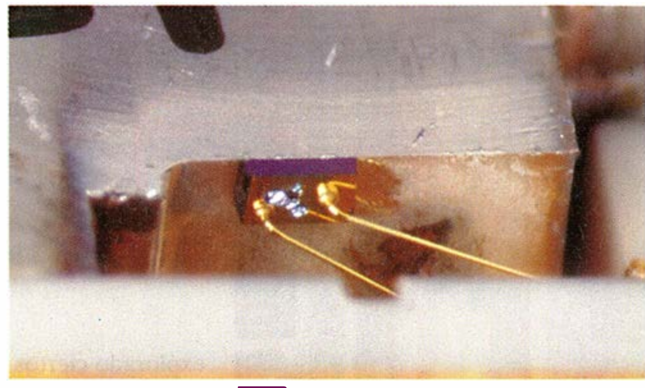
Los materiales en los que se pro-

DOS DIMENSIONES: EL POZO CUANTICO



La incorporación de un pozo cuántico sencillo (*izquierda*) en este láser de diodo (*derecha*) mejora sustancialmente el rendimiento de dicho láser. Las capas exteriores contaminadas inyectan electrones (cargas negativas) y agujeros (cargas positivas) en el pozo cuántico; estas partículas se recombinan posteriormente y producen una emisión de luz. Porque el pozo cuántico contiene menos aluminio que las capas que lo rodean, sabemos que se trata de un material cuyo intervalo entre bandas es menor.





ARSENIURO DE GALIO

ARSENIURO DE GALIO Y ALUMINIO

ARSENIURO DE GALIO Y ALUMINIO CONTAMINADO CON CARGAS NEGATIVAS

ARSENIURO DE GALIO Y ALUMINIO CONTAMINADO CON CARGAS POSITIVAS

duce de manera natural este confinamiento de los electrones han intrigado durante mucho tiempo a los científicos. El grafito, un material orgánico conductor de la electricidad, está constituido por una superposición de capas bidimensionales de átomos de carbono. Algunas cerámicas superconductoras a temperaturas altas, tales como el óxido de cobre y lantano, están formadas por planos bidimensionales de átomos de cobre y oxígeno separados por planos formados por otros átomos.

Para construir deliberadamente un material con menos dimensiones que las normales, hemos de basarnos en la mecánica cuántica. En cualquier material semiconductor ordinario tridimensional, los electrones se sitúan en alguno de los posibles estados energéticos de un intervalo continuo cuando se suministra cierta energía adicional al material en el que se encuentran; por ejemplo, aplicando un voltaje. Los investigadores no pueden, pues, acceder a un determinado estado energético, sino que se ven obligados a aceptar lo que se les ofrece.

La teoría mecánico-cuántica considera a los electrones como partículas y ondas. Cuando se reduce la longitud de una de las caras de un cubo tridimensional, hasta conseguir que dicho espesor no sea mayor que la longitud de onda de los electrones, éstos quedan atrapados en un plano bidimensional. En tal espacio de dos dimensiones, la densidad de estados electrónicos —los niveles de energía que los electrones pueden ocupar— queda cuantizada. Los electrones saltan desde un nivel de energía hacia otro en forma discreta. Estudiando cuál es el espesor de la capa que induce un nivel de energía concreto, se pueden diseñar las características electrónicas precisas de un material.

Pero los electrones en cuestión no están realmente confinados por barreras físicas; antes bien, a los investigadores cumple levantar barreras energéticas. De forma parecida a lo que sucede cuando el agua baja por los torrentes, los electrones tienden a ocupar las zonas con energías más bajas. Por tanto, para atrapar electrones, sólo se necesita colocar una lámina de material con electrones de energía baja entre dos láminas de materiales con electrones de energía elevada.

La energía de los electrones se puede describir mediante la teoría de bandas. Los electrones de valencia, o electrones exteriores de cada átomo, se sitúan en unas bandas o intervalos de energía específicos. Los átomos metálicos poseen unos pocos electrones de valencia con energías bajas. Dado que los átomos metálicos intercambian sus electrones con mucha facilidad, los metales son buenos conductores de la electricidad.

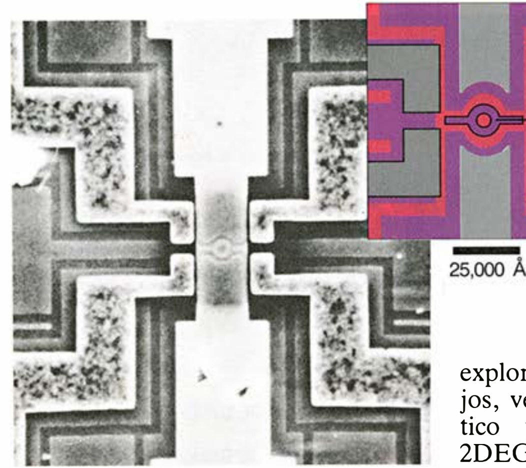
Por otra parte, los semiconductores y aislantes poseen más electrones de valencia y no son buenos conductores de electricidad. Cuando suministramos una pequeña cantidad de energía a un semiconductor, algunos electrones de valencia se transfieren hasta un nivel de energía más elevada: hasta la banda de conducción; ello permite que el material conduzca la corriente eléctrica. La cantidad de energía necesaria para transferir un electrón de una a otra banda recibe el nombre de intervalo de energía o "gap" entre las bandas; representa la diferencia entre los niveles energéticos de la banda de valencia y la banda de conducción. Dado que la cantidad de energía necesaria para hacer posible la conducción difiere ligeramente de un semiconductor a otro, el intervalo de energía de un determinado semiconductor es mayor o menor que el de los demás. Los aislantes requieren una enorme cantidad de energía para transferir sus electrones de valencia hasta las bandas de conducción de energía más alta; en este caso, el intervalo entre bandas resulta especialmente grande.

Los investigadores de las empresas IBM y AT&T intentaron demostrar la imposibilidad del confinamiento de los electrones. Esaki, Tsu y Chang empezaron este trabajo apilando una serie de capas alternas de arseniuro de galio y arseniuro de aluminio y galio, un compuesto cuyo intervalo entre las bandas es más elevado que el del primero. El equipo de AT&T se fijó un objetivo mucho más sencillo: colocar un material cuyo intervalo entre bandas era pequeño o entre dos materiales cuyo intervalo entre bandas era mayor, produciendo un pozo cuántico.

Los esfuerzos de ambos grupos se vieron dificultados por muchos pro-

UNA DIMENSION: EL CABLE CUANTICO

Una capa metálica transparente cubre este bucle de cable fabricado en IBM, que se emplea para realizar mediciones eléctricas en un campo magnético (*izquierda*). Los cables cuánticos fabricados en Bellcore se forman en la parte más interna de una muesca en V practicada en un material semiconductor (*derecha*).



blemas de fabricación. ¿Cómo apilar capas de materiales con espesores de tan sólo algunos átomos? Chang recuerda que, para conseguir depositar estas capas tan delgadas, "tuvimos que construir nosotros mismos un sistema de vacío adecuado". Otro problema importante consistió en controlar la contaminación de los sustratos y la alineación de las redes cristalinas de los materiales.

A pesar de todas las dificultades llegó el éxito en 1974. El grupo de IBM obtuvo materiales que poseían los niveles energéticos discretos que predecía el confinamiento cuántico. Raymond Dingle y sus colaboradores, de los Laboratorios Bell, construyeron un único pozo cuántico, hicieron incidir luz láser sobre él y observaron que el nuevo material absorbía la luz cuyas frecuencias discretas coincidían con las esperadas, nueva señal de que se había producido el confinamiento cuántico. Poco después, Esaki y Chang construyeron un diodo de efecto túnel resonante, que fue el primer dispositivo real constituido por un pozo cuántico.

Transistores balísticos

Otros investigadores escogieron una vía diferente, partiendo de los llamados gases electrónicos bidimensionales (conocidos por las siglas 2DEG), en los que los electrones resultan atrapados en una interfase horizontal situada entre una capa de material de intervalo energético bajo y una capa contaminada con cargas adicionales. Con estos materiales se han explorado ciertos fenómenos complejos, verbigracia, el efecto Hall cuántico y fraccional. En materiales 2DEG construidos con sumo cuidado, los electrones recorren tramos considerables antes de sufrir una colisión con un átomo o un defecto; merced a ello, los ingenieros pueden construir transistores "balísticos" o con "alta movilidad de los electrones".

Los 2DEG constituyen la piedra angular de los transistores de alta velocidad. (Estos dispositivos sólo alcanzan las máximas velocidades cuando trabajan a temperaturas de grados Kelvin.) Al contrario de lo que sucede con los pozos cuánticos, en los que los electrones se ven obligados a atravesar varias capas de material dispuestas verticalmente, los electrones de los 2DEG se mantienen dentro de un plano horizontal. Esta particularidad restringe los tipos de dispositivos que se pueden construir con ellos.

A principios de los años ochenta, los investigadores volvieron a los pozos cuánticos y a las superredes, que no son otra cosa que pozos cuánticos múltiples construidos especialmente, y esta vez lo hicieron armados con herramientas de fabricación muy precisas. Entre esos nuevos útiles citaremos el empleo de máquinas generadoras de haces moleculares de capas delgadas, conocidos por las siglas MBE.

El elemento fundamental de las modernas máquinas MBE es una cámara de vacío ultraelevado, que deja que los investigadores depositen capas de átomos con un espesor del orden de dos angstroms sobre un sustrato de material semiconductor. Esta cámara de vacío se halla conectada con tres o cuatro conductos que la unen con las células de efusión. Los elementos como el galio o el aluminio se vaporizan en estas células y se envían a través de dichos conductos hacia el sustrato base. Se controla el espesor de las capas de material depositadas manipulando adecuadamente las compuertas que separan los conductos y la cámara de vacío. Se necesita todo un día de trabajo para depositar una capa de semiconductor de unas ocho micras de espesor sobre un único sustrato de arseniuro de galio.

Sin embargo, en comparación con la fase siguiente del proceso, el crecimiento de estas capas es sencillo. Para hacer uso de los pozos cuánticos, los investigadores deben labrar o esculpir instrumentos —láseres por ejemplo— de esas capas apiladas. (Sirven de láseres los semiconductores emisores de luz cuando se excitan con luz o electricidad.)

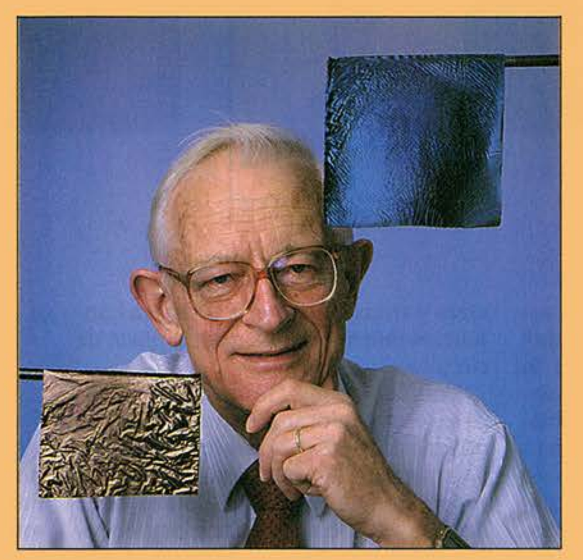
Para conseguir este propósito, uno de los investigadores de Bellcore, Edward M. Clausen, Jr., ha construido una máquina grabadora que emplea haces iónicos con asistencia química. Mediante un proceso litográfico, Clausen graba una figura sobre el sustrato. Introduce luego dicho sustrato en una cámara llena de gas cloro con el fin de acelerar el ataque químico y lo bombardea con iones (átomos cargados). En teoría, los iones atravie-

POLIMEROS CONDUCTORES EN UNA DIMENSION

En el año 1975, Alan G. MacDiarmid, químico de la Universidad de Pennsylvania, acababa de pronunciar una conferencia en el Instituto de Tecnología de Tokyo. Sus anfitriones le invitaron a visitar las distintas dependencias del centro. Cuando se encontraba en el laboratorio de Hideki Shirakawa, vio una película plateada que le gustó mucho. "Nunca había visto un polímero plateado", confiesa McDiarmid.

Los químicos que trabajaban con polímeros buscaban sistemas electrónicos de una dimensión, pero se encontraron con ellos en las láminas de polímero de Shirakawa. Los materiales resultaron ser una nueva forma de poliacetileno, un polímero simple cuyo esqueleto está constituido por una serie de átomos de carbono, con un átomo de hidrógeno unido a cada átomo de carbono. Como la mayoría de compuestos orgánicos, el poliacetileno es un aislante; sus átomos se unen entre sí por enlaces covalentes sencillos y todos sus electrones de valencia están compartidos, lo cual no permite la presencia de electrones adicionales capaces de transportar una corriente eléctrica. Sin embargo, al sintetizar este material, Shirakawa había creado accidentalmente una película polimérica con una estructura especial de enlaces entre los átomos, consistente en una sucesión de enlaces simples y dobles alternativos. Resultaba muy útil, pues permitía el paso de electrones a través del material.

Trabajando juntos, MacDiarmid, Shirakawa y Alan J. Heeger se aprestaron a contaminar las películas de polímero, añadiendo o extrayendo electrones. Con ello crearon nuevas bandas de energía entre las bandas de conducción y de valencia. Las películas plateadas tomaron entonces un aspecto dorado y se convirtieron en conductoras.



ALAN G. MACDIARMID desarrolló los polímeros conductores. La polianilina contaminada (derecha) *conduce la corriente eléctrica; la muestra sin contaminar* (izquierda) *no lo hace.*

Se han descubierto otras familias de polímeros conductores, la polianilina entre ellos, que incluye átomos de nitrógeno en la cadena del esqueleto básico, y el politiofeno, que cuenta con azufre.

PINZAS MICROSCOPICAS. La característica física de todos estos polímeros, la conductividad a una densidad excepcionalmente baja, provocó en un principio muchas expectativas. La fabricación de fuselajes de aviones y baterías para automóviles con muy poco peso se encontraban entre las primeras propuestas. La contaminación y no contaminación de estos polímeros modifica también su longitud hasta en un 10 por ciento. En virtud de esta propiedad, Raymond Baughman, de la firma Allied Signal, sugirió el empleo de polímeros conductores para convertir energía eléctrica en energía mecánica.

A pesar de todo, la plena utilización de todas las posibilidades electrónicas que ofrece la estructura casi unidimensional de estos polímeros ha requerido mucho tiempo. La ciencia se pregunta todavía el motivo por el que estos polímeros conducen la electricidad y cómo mejorar su conductividad. Lo que sabe actualmente es que los electrones se mueven a lo largo de la cadena de un polímero –como si se tratara de un cable unidimensional– hasta que se encuentran con un defecto o una rotura de la cadena. Dado que los electrones se comportan como ondas, rebotan en la barrera, retornando por el mismo camino por donde habían llegado. Cuando los electrones rebotan en un sentido y otro entre dos defectos, se produce una onda estacionaria. Esta situación origina una retención de la carga correspondiente en una zona restringida del polímero y su conductividad general disminuye.

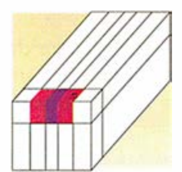
SALTANDO LA CADENA. Los electrones pueden escapar de esta trampa, aumentando espectacularmente la conductividad del material, si "saltan" a una cadena del polímero próxima antes de que se forme la onda estacionaria. (La posibilidad de que se produzca este salto lateral es lo que convierte a los polímeros en sistemas casi unidimensionales.) Este salto de los electrones puede ser muy parecido al efecto túnel que experimentan los electrones cuando se encuentran confinados en un semiconductor. La conductividad general de un polímero dependerá de la distancia que pueden recorrer los electrones antes de encontrarse con algún defecto en la red, así como de su facilidad para saltar de una cadena a otra. Ambos factores están relacionados con la concentración de defectos a lo largo de las cadenas.

Algunos polímeros muestran interesantes características ópticas. Pueden transmitir la luz procedente de un láser. Y como consecuencia de la forma en que los polímeros comparten electrones, un aumento en la intensidad de la luz producirá una modificación en las propiedades ópticas del material en cuestión.

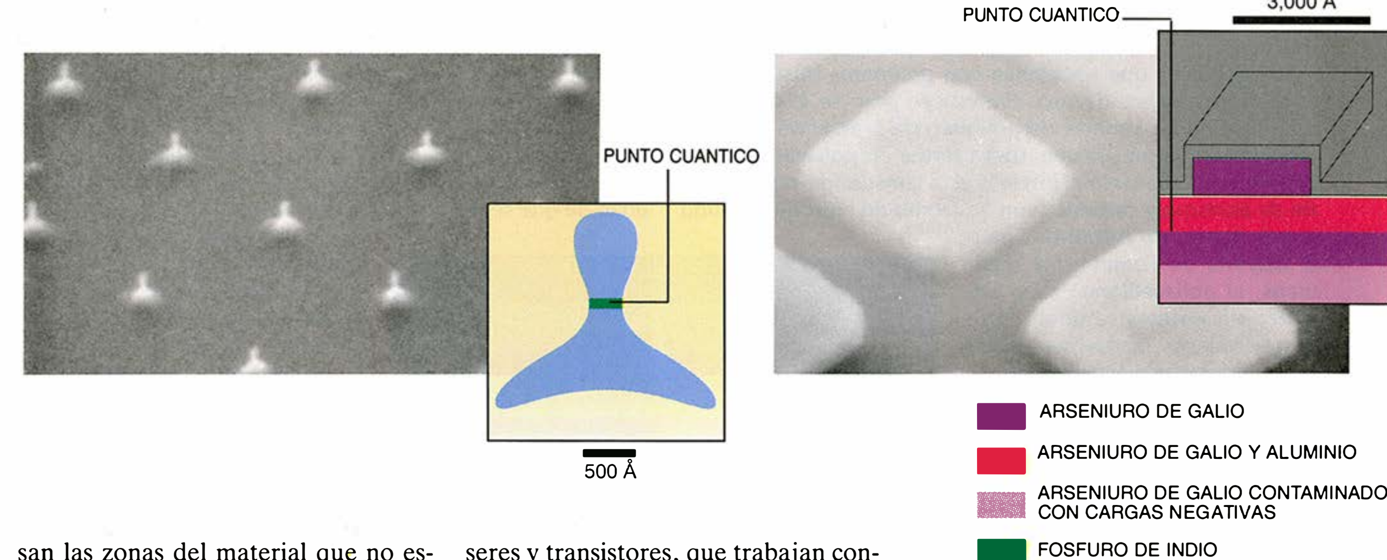
CONMUTADORES OPTICOS. Se confía en aprovechar estas propiedades para fabricar conmutadores ópticos, es decir, dispositivos que utilicen luz para controlar la luz. Shahab Etemad, de Bellcore, está desarrollando uno de estos conmutadores poliméricos, basado en dos filamentos idénticos de polidiacetileno que discurren paralelos en el seno de un bloque de vidrio. Estos filamentos, cuyo diámetro es menor que la longitud de onda de la luz que transmiten, son en realidad guías de ondas acopladas. Como resultado, cuando se transmite la luz de un láser de baja intensidad a lo largo de uno de los filamentos, se produce una oscilación de la luz entre ambas líneas.

Para convertir a las guías de ondas en un conmutador, Etemad procede a aumentar la intensidad de la luz de láser, lo que produce un cambio en el índice de refracción de la línea superior. Este hecho da lugar a que las dos líneas dejen de ser idénticas y, por tanto, las dos guías de onda se desacoplan. Controlando la longitud de estas líneas, la luz de alta intensidad sale sólo a través de la línea superior y la luz de baja intensidad lo hace sólo a través de la línea inferior. El tiempo necesario para conectar (o desconectar) un filamento polimérico es de un femtosegundo (es decir, 10^{-15} segundos). Esta velocidad es mucho mayor que la que permite cualquier conmutador de semiconductores.

CERO DIMENSIONES: EL PUNTO CUANTICO



En AT&T se fabrican puntos cuánticos en un proceso al vacío (*figura de la izquierda, en esta página*). Los puntos definidos litográficamente y construidos por IBM sirven de condensadores. En estos dispositivos, los electrones quedan confinados en la interfase existente entre sendas capas de arseniuro de aluminio y galio y arseniuro de galio (*derecha, en esta página*). Las matrices de puntos fabricadas por Texas Instruments (*izquierda, en la página siguiente*) se pueden utilizar como transistores. Los químicos fabrican estos puntos capturando diferentes agregados de átomos (cadmio, en color naranja, y azufre, en color amarillo) en "jaulas" de zeolita, según se muestra en esta simulación obtenida con ordenador por Galen D. Stucky en la Universidad de California en Santa Bárbara (*derecha, en la página siguiente*).



san las zonas del material que no están recubiertas por la máscara y dejan intactas las zonas protegidas. En la práctica, los iones pueden dañar también el sustrato de forma irreversible.

Clausen y sus colaboradores han dedicado mucho tiempo al tallado de láseres microscópicos a partir de pozos cuánticos múltiples. Esfuerzos que están empezando a rendir sus frutos. A principios de 1990, un equipo conjunto de Bellcore y AT&T ocupó los titulares de los periódicos después de conseguir grabar casi dos millones de láseres en una superficie algo menor que un centímetro cuadrado. Al contrario de lo que sucede en los láseres de semiconductores convencionales, que emiten luz por uno de sus bordes, la emisión de estos nuevos láseres tiene lugar por la superficie superior o inferior.

La incorporación de un pozo cuántico en un láser presenta una ventaja muy importante: este dispositivo puede producir luz con un rendimiento muy elevado, consumiendo mucha menos energía que la que se necesita en los láseres al uso. Por ello, el calor emitido por los láseres cuánticos es mucho menor. Esta cualidad, unida al pequeño tamaño de estos láseres, permite el empaquetamiento compacto de los dispositivos. De esta forma, los emisores superficiales que incorporan pozos cuánticos pueden convertirse en las piezas fundamentales de los modernos dispositivos optoelectrónicos; estos dispositivos están constituidos por la unión de láseres y transistores, que trabajan conjuntamente sobre una misma placa de circuito.

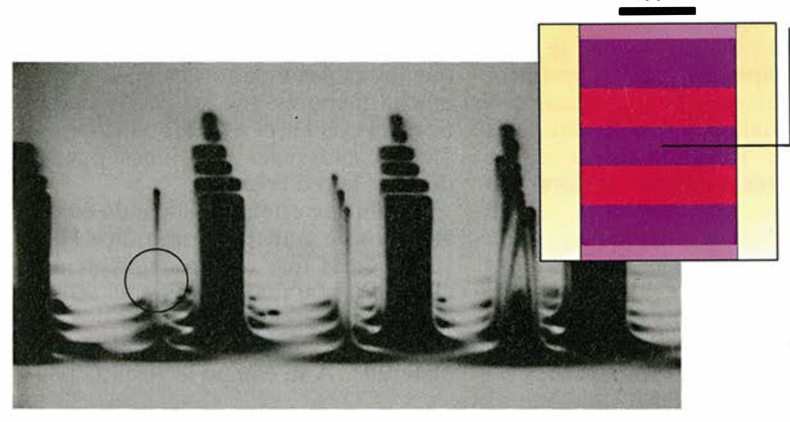
Connie Chang-Hasnain, de Bellcore, está explorando otras formas de utilización de estos láseres en comunicaciones. Tiene sobre su mesa de laboratorio una fotografía de una serie de columnas constituidas por pequeños láseres de pozos cuánticos. Cada uno de ellos está rodeado por una almohadilla metálica de contacto. Unas sondas metálicas afiladas, con una forma parecida a la de mosquitos gigantes, se apoyan sobre tres de las almohadillas de contacto de los láseres. Con grandes precauciones, Chang-Hasnain hace llegar una corriente de algunos miliampères hasta las sondas. El resultado se pone de manifiesto en otra pantalla de ordenador, sobre la que aparecen tres puntos luminosos separados.

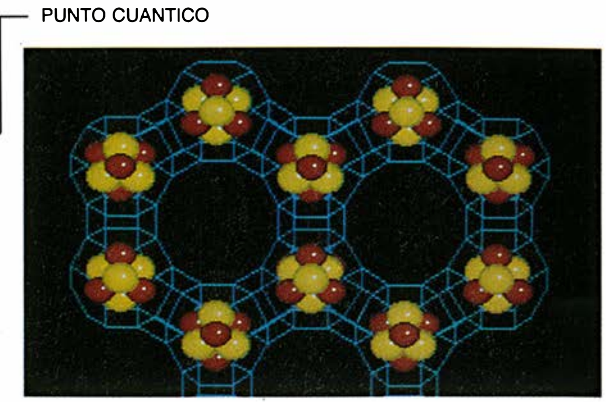
Los datos de Chang-Hasnain ponen de manifiesto que cada uno de los láseres emite una frecuencia luminosa diferente. Controlando cuidadosamente la rotación del sustrato semiconductor durante el crecimiento de la MBE, esta investigadora y sus colaboradores han hecho crecer un bosque microscópico de láseres. Debido a que cada dispositivo tiene capas de espesores ligeramente diferentes, los láseres son únicos. El mejor resultado obtenido hasta ahora lo cifran en un conjunto de 77 láseres, cada uno de los cuales es capaz de transmitir datos con su propia frecuencia. La conexión de este conjunto de láseres a una línea de fibra óptica "sería —apostillan— como añadir miles de canales más a su televisión por cable".

Paul L. Gourley, de los Laboratorios Sandia en Albuquerque, objeta que la unión de componentes ópticos y electrónicos presenta todavía muchas dificultades. Gourley está construyendo también conjuntos de láseres de superficies emisoras; se muestra de acuerdo en que estos dispositivos pueden ofrecer un elevado rendimiento. Pero señala que el acoplamiento de láseres muy pequeños a fibras ópticas exige una finísima precisión. Por otra parte, dado que las respectivas ondas luminosas se mantienen en fase, o tienen una longitud de coherencia mayor que las ondas asociadas de los electrones, la construcción de estos componentes optoelectrónicos integrados puede resultar muy estimulante.

Dimensiones prácticas

Otros autores continúan expresando su preocupación acerca de los enormes problemas que plantea la fabricación de estos láseres. La transición de estos prototipos de laboratorio a sistemas producidos en serie exigirá todavía mucho trabajo. Tras más de 15 años de investigación, sólo han





llegado al mercado algunos de estos láseres cuánticos. Uno de estos dispositivos consiste en un láser que emite por los bordes y que consta de un pozo cuántico; este láser ha sido fabricado por la empresa Spectra Diode Laboratories. El dispositivo posee una potencia de 100 miliwatts y puede convertir en luz hasta un 60 por ciento de la energía eléctrica que lo alimenta, lo cual supone una eficiencia doble que la mayoría de los láseres de semiconductores de características comparables. (La unión de varios de estos láseres produce una emisión de luz mayor, pero con un menor rendimiento.) Sin embargo, estos láseres deben construirse de uno en uno.

John N. Randall, de Texas Instruments, dedica también la mayor parte de su tiempo a grabar sobre capas de pozos cuánticos. Randall desea construir transistores.

Hasta ahora, los fabricantes de circuitos integrados han tenido éxito en la constante miniaturización de pastillas, transistores o conmutadores. Pero cuando las dimensiones de los circuitos se acercan al tamaño de la longitud de onda de los electrones, éstos tienden a escaparse del dispositivo, impidiendo que los transistores puedan desconectar realmente el circuito. Robert T. Bate, responsable del trabajo cuántico en la empresa Texas Instruments, apostilla: "Nosotros ya preveíamos todos estos problemas prácticos a finales de la década", cuando los fabricantes de pastillas querían circuitos con un espesor inferior a 0,25 micras, o sea, 2500 angstroms.

Randall y sus colaboradores esperan salvar estos escollos recurriendo a los efectos cuánticos para construir transistores mucho más pequeños que los fatídicos 2500 angstroms. Con tal fin, están desarrollando descendientes refinados de los primitivos diodos resonantes de efecto túnel. Un éxito notable en esta línea lo constituye el transistor bipolar cuántico resonante de efecto túnel, conocido por sus siglas en inglés, BiQuaRTT.

Según explica Reed, diseñador de los BiQuaRTT, todas las variedades de diodos y transistores resonantes de efecto túnel obedecen a las mismas reglas básicas. Como punto de partida, se toma una capa delgada de pozo cuántico colocada entre dos láminas igualmente delgadas pero cuyos intervalos entre bandas son mayores. A continuación, se aplica un voltaje. Para muchos voltajes, los electrones no pueden atravesar las capas del dispositivo, debido a las elevadas barreras que oponen los intervalos entre bandas, y el dispositivo conduce poca corriente.

Sin embargo, de acuerdo con la mecánica cuántica, existe cierta probabilidad de que los electrones atraviesen dichas barreras en virtud del denominado "efecto túnel". Las consecuencias de este efecto túnel se pueden reforzar aplicando el voltaje adecuado al dispositivo. Para estos voltajes específicos, la energía de los electrones incidentes coincide con los niveles energéticos del pozo cuántico. En este caso, los electrones atravesarán la barrera como si pasaran a través de un túnel y se producirá el tránsito de corriente.

Como es natural, un circuito integrado no está constituido por un solo transistor. (Algunos circuitos integrados están formados por más de un millón de transistores.) En Fujitsu se han conectado hasta cinco o seis transistores resonantes de efecto túnel en serie con conmutadores de alta velocidad. Texas Instruments espera conseguir en un plazo máximo de 30 meses que los BiQuaRTT manejen lógica compleja y, así, poder construir circuitos que sean varios órdenes de magnitud más pequeños que los actuales. Dado que el efecto túnel se produce realmente para dos voltajes diferentes, "estos dispositivos —expone Randall— nos permitirán construir conmutadores para dos voltajes diferentes". Espera también que estos conmutadores sustituyan a ciertos transistores tradicionales y que varios BiQuaRTT trabajen conjuntados.

Si las perspectivas bidimensionales son prometedoras, las de una o cero dimensiones son todavía mejores. Los electrones se difunden menos cuando se mueven a través de cables conductores cuánticos, alcanzando movilidades más altas y facilitando la construcción de conmutadores más rápidos. Los láseres de cables conductores cuánticos podrían alimentarse con corrientes mucho más débiles y radiarían, por tanto, una cantidad de calor mucho menor que sus equivalentes bidimensionales.

Cables de semiconductores

Ya hemos visto bastantes teorías. Para construir cables cuánticos, los investigadores deben recubrir los cuatro lados de un material cuyo intervalo entre bandas sea pequeño con un material de barrera con un intervalo entre bandas grande y con un espesor lo suficientemente pequeño como para permitir que los electrones puedan atravesarlo por efecto túnel cuando se desee; este espesor debe ser del orden de la longitud de onda de dichos electrones. El control del proceso para construir estas barreras verticales es sumamente difícil, dificultad que ha llevado a muchos investigadores a reestudiar la forma en que se fabrican las capas de semiconductores.

En el laboratorio de investigación tecnológica optoelectrónica en la Ciudad de la Ciencia Tsukuba, con sólo cuatro años de operación, Toshiro Isu observa una vez más su cinta de vídeo en blanco y negro, rayada ya por el uso. Desde enero del año pasado, Isu ha venido grabando en imágenes el

crecimiento de las capas de galio y arsénico en una cámara de vacío MBE. Las imágenes recogen la formación de conjuntos constituidos por millones de átomos de galio en breves segundos; estos agregados parecen fundirse, a continuación, en el seno de las capas de arsénico preexistentes. A Isu le gustaría saber por qué dichos conjuntos aparecen en lugares aleatorios y cuál es su efecto sobre la uniformidad de las capas del cristal.

Izuo Hayashi, del mismo laboratorio y uno de los inventores del láser de semiconductores, reconoce que nadie sabe cómo tiene lugar la formación de capas delgadas por haces moleculares y de qué manera se combinan las capas de galio y arsénico; antes de que se puedan fabricar dispositivos altamente confinados tales como cables y puntos cuánticos, los investigadores deberán mejorar su conocimiento acerca de lo que sucede a nivel atómico durante el proceso de fabricación y grabado de las capas delgadas.

A pesar de ello, los esfuerzos para fabricar cables y puntos cuánticos continúan. Muchos investigadores están ensayando las mismas técnicas que se utilizan en la construcción de pozos cuánticos. Sin embargo, dado que los cables y puntos cuánticos son mucho menores que los pozos cuánticos, el bombardeo de los correspondientes materiales con iones puede destruir la red cristalina.

De ahí que estén investigando otros métodos de trabajo. Como dice Hayashi, "más que cortar simplemente el material (grabar a través del semiconductor), necesitamos hacer uso de alguna tendencia natural de dicho material". A su equipo le gustaría po-

ELECTRONES QUE SE MUEVEN DE UNO EN UNO

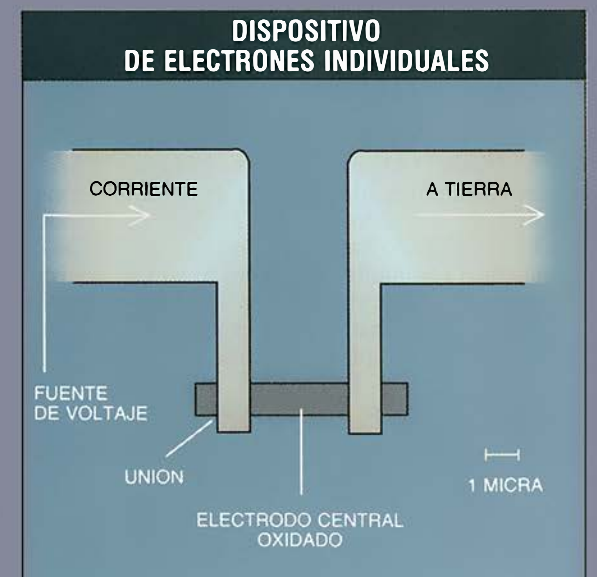
La corriente eléctrica que atraviesa los dispositivos electrónicos convencionales se puede asimilar a un río, donde los electrones abundarían como las moléculas del agua. Cuando se reducen las dimensiones de estos materiales, la energía de dichos electrones resulta cuantizada, es decir, dividida en incrementos discretos. A pesar de ello, el número de electrones presentes en estos materiales supera todos los cálculos.

Los investigadores de la Universidad de Delft, el Centro de Estudios Nucleares de Saclay, la Politécnica de Chalmers y la Universidad de Moscú acaban de demostrar que es posible contar el número de electrones que contribuyen a una corriente poco intensa. En vez de la energía cuantizada de los electrones, estudian las cargas electrónicas cuantizadas.

Las implicaciones de tanto esfuerzo son dobles: los dispositivos utilizados para contar electrones uno a uno pueden conducir a una nueva norma de gran precisión para medir la intensidad de la corriente. A más largo plazo, las corrientes eléctricas formadas por cierto número de electrones individuales podrían conducir a esquemas alternativos para construir dispositivos electrónicos a muy pequeña escala.

ELECTRONES SINCOPADOS. El diseño básico de un dispositivo de un solo electrón consta de un electrodo central de aluminio con una longitud del orden de una micra, recubierto con óxido de aluminio. Esta barra está montada sobre dos electrodos metálicos dispuestos simétricamente en cada uno de sus extremos. Aunque sus dimensiones son muy pequeñas, estos dispositivos no constituyen, hablando en sentido estricto, sistemas dimensionalmente confinados.

Cuando se aplica un voltaje a través de este dispositivo, se comporta como un resistor convencional. Mas, para temperaturas del orden de un grado Kelvin y voltajes de algunas décimas de milivolt, su resistencia eléctrica experimenta un incremento extraordinario. Ninguna corriente logrará atravesar este material porque la energía que poseen los electrones es insuficiente para cargar la barra central. Cuando el voltaje aumenta hasta alcanzar un valor de milivolt, los electrones adquieren la energía necesaria para pasar por efecto túnel desde el electrodo al que se le ha aplicado el voltaje hasta el electrodo central.



Através de la barra del dispositivo sólo puede pasar un electrón a un tiempo. La explicación de este hecho radica en que la llegada de un único electrón carga la barra y eleva así la barrera de energía que el siguiente electrón tendría que superar para llegar a esa misma barra. Cuando el primer electrón abandona la barra, la carga de la barra disminuye. En consecuencia, el flujo de los electrones transcurre de forma altamente sincopada: es necesario que un electrón abandone la barra central antes de que otro electrón entre en ella.

UNO ENTRE MIL. Para utilizar este efecto de electrón individual como norma precisa de intensidad de corriente, los investigadores de Delft y Saclay sincronizaron la corriente producida por un solo electrón con una señal de radiofrecuencia. Empezaron por construir una serie de pequeños electrodos. Uniendo fuentes de voltaje adicionales a los electrodos, las barreras de energía de este dispositivo pueden elevarse o reducirse a voluntad. A continuación, se acopla una fuente secundaria de voltaje a una señal de radiofrecuencia, para conseguir que esta señal regule el paso de electrones a través del dispositivo, a razón de uno por ciclo.

Hasta ahora, los investigadores han registrado el paso de corrientes de hasta algunos picoampères, medidas con una precisión del orden del femtoampère, para frecuencias comprendidas entre los 5 y los 20 millones de electrodos por segundo. Por tanto, estas medidas son precisas hasta el orden de 1 electrón cada 1000. La aplicación de este método permitiría fijar la norma establecida para la intensidad de la corriente eléctrica con una precisión superior a 1 en 10 millones.

Las puertas para electrones construidas en Delft y Saclay trabajan de forma similar a como lo hace un registrador de desplazamiento ordinario, un dispositivo electrónico destinado al almacenamiento de números binarios. Por otra parte, los dispositivos de electrón individual pueden resultar más fáciles de fabricar que los alambres y puntos cuánticos; la construcción de puertas metálicas no se basa en uniones complejas de materiales compuestos diferentes.

Hasta ahora, la investigación se ha realizado a temperaturas próximas a un grado Kelvin. Está por ver si los efectos relacionados con los electrones individuales se mantendrán cuando la temperatura aumente.

der controlar los bordes de las capas atómicas delgadas a medida que el material va creciendo en un MBE.

Otros grupos están ensayando esquemas diferentes. Pierre M. Petroff, Arthur Gossard y sus colaboradores, de la Universidad de California en Santa Bárbara, desarrollan cables cuánticos en un MBE apilando elementos, de capa atómica en capa atómica, sobre una serie de escalones dispuestos en forma de terrazas de un sustrato semiconductor. Esta técnica permite obtener pozos cuánticos verticales (o más frecuentemente, inclinados) cuyo lado mide 50 angstroms y que pueden convertirse en cables cuánticos. Sin embargo, para conseguir esto se necesita una extraordinaria precisión. Ellos mismos se propusieron luego cambiar de forma continua el momento en que las capas atómicas crecen sobre la superficie en que están depositadas, de suerte que las capas resultantes no se disponen creando una línea recta o inclinada, sino que muestran una disposición ondulada. En esta "superred serpentina", los cables cuánticos se forman en las zonas curvadas, y los correspondientes conjuntos de cables son mucho más uniformes que los obtenidos con las técnicas utilizadas antes.

Eli Kapon y sus colaboradores de Bellcore optaron por otras técnicas de grabado. Practicaron una estría en V sobre un sustrato de material semiconductor y, a continuación, depositaron en la base de dicha V un material con un pequeño intervalo entre bandas. El equipo de Kapon ha construido láseres constituidos por uno, dos o tres cables cuánticos. Estos láseres producen una emisión de luz incluso cuando se alimentan con una corriente de 0,65 miliampères a temperatura ambiente, un valor umbral pequeñísimo para un láser.

B. E. Maile y sus colaboradores, de la Universidad de Stuttgart, fabricaron un conjunto de cables sobre un sustrato utilizando haces iónicos reactivos. Limpiaron luego los cables con sumo cuidado mediante un disolvente químico y, finalmente, los "enterraron" cubriéndolos con una capa de material semiconductor utilizando una técnica de deposición similar a la MBE, consistente en la formación de capas delgadas de materiales organometálicos en fase de vapor (un proceso conocido por las siglas en inglés, MOVPE).

Se sigue discutiendo si todas estas técnicas, o cualesquiera otras, han sido capaces de proporcionar realmente el desdoblamiento singular de las energías de los electrones que caracterizan a los cables cuánticos. La mayoría de estos cables son demasiado gruesos. Dado que el desdoblamiento de energía resulta ser menor cuanto mayor es el tamaño de los correspondientes dispositivos, cuesta afirmar que estamos ante cables cuánticos.

La zona cero

Al final del arco iris cuántico se encuentran los puntos cuánticos. Estos elementos responden también a la denominación de átomos artificiales, aunque pueden estar formados por miles o centenares de miles de átomos. Los electrones ocupan estados energéticos discretos cuando están confinados en un punto o caja. En consecuencia, la obtención de unos niveles energéticos precisos debería ser factible fabricando una caja cuántica adecuada y modificando el voltaje aplicado.

Yasuhiko Arakawa e Hiroyuki Sakaki propusieron el concepto de cajas o puntos cuánticos en el año 1982. No han podido construir, sin embargo, ningún dispositivo de ésos que funcione. Confiesa el propio Arakawa: "Es muy difícil construir un pozo cuántico cuyas dimensiones (a lo largo de sus seis lados) sean del orden de 100 angstroms." Y añade: "Nadie ha conseguido hasta ahora fabricar un láser de pozo cuántico."

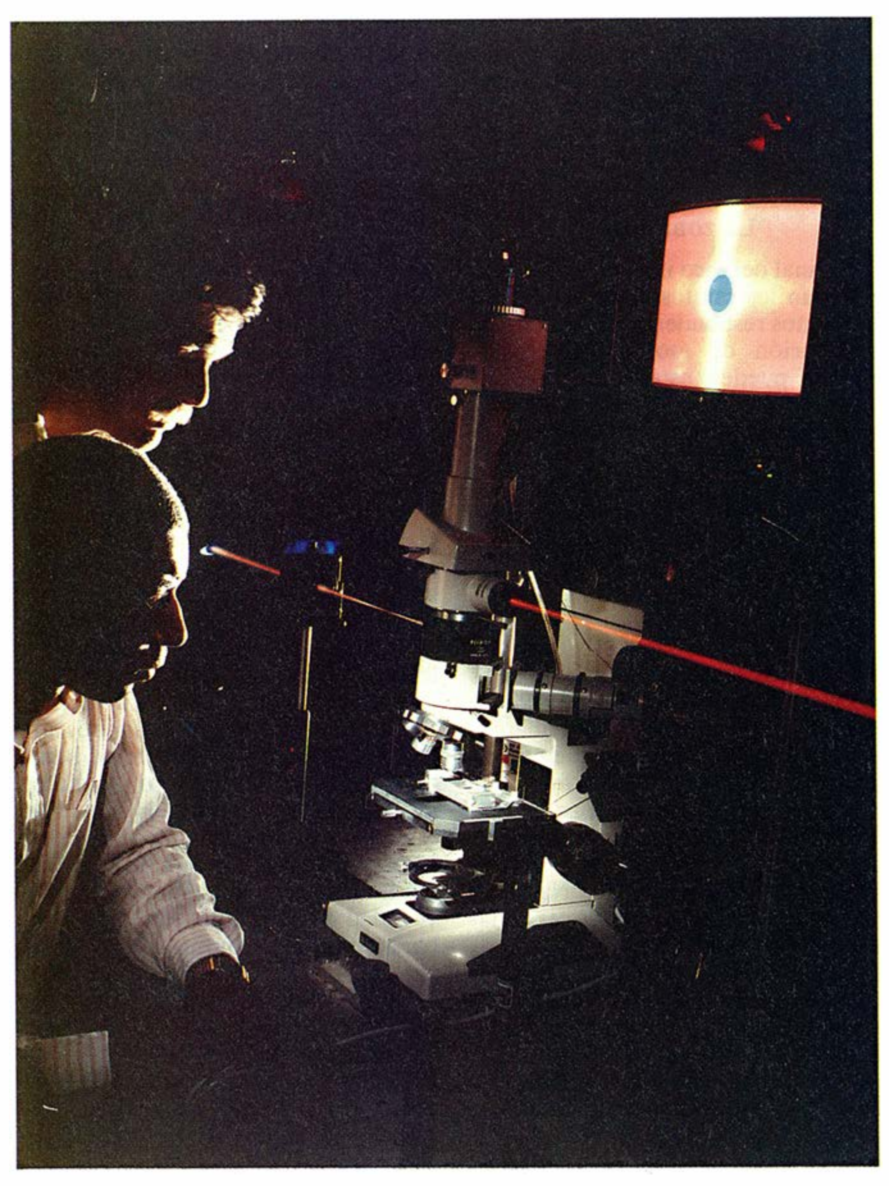
La fabricación de estos puntos cuánticos utilizando la técnica de los semiconductores parece presentar dificultades insuperables. La afirmación de que uno de estos puntos "trabaja" realmente es siempre objeto de controversia. Y un único punto en sí mismo no reviste particular interés. Hay que encontrar, además, la forma de fabricar colecciones de puntos y, a continuación, integrarlas en los oportunos dispositivos. Los químicos, que han logrado muchos éxitos en la fabricación de puntos cuánticos, no han conseguido sean implantables en tales dispositivos. En su despacho de los Laboratorios Bell, Michael L. Steigerwald posee una colección de frasquitos. Cada uno de ellos está lleno con un polvo de vivos colores: amarillo, naranja, rojo, negro. "Todos estos productos son seleniuro de cadmio", señala. Cada frasco contiene cierta cantidad de partículas de tamaños diferentes, constituidas por átomos de cadmio y selenio, y que no son otra cosa que puntos cuánticos.

Las partículas de tamaños diferentes tienen propiedades físicas singulares, tales como una dureza excepcional, y absorben unas determinadas frecuencias luminosas. Por ejemplo, el polvo amarillo está constituido por partículas de seleniuro de cadmio cuyo tamaño es del orden de 15 angstroms, es decir, cada una de estas partículas contiene unos 100 átomos. Las partículas de color rojo miden unos 40 angstroms y están formadas por unos 1200 átomos. La frecuencia de la luz absorbida depende de la separación entre bandas del material que constituye la partícula. Como explica Steigerwald, cuanto menor es el tamaño de una partícula, tanto mayor resulta ser la separación entre bandas. En consecuencia, los niveles de energía de un material se pueden ajustar haciendo variar el tamaño de las partículas.

Para fabricar estas partículas, los químicos encierran pequeños grupos de átomos en unas vainas o jaulas especiales. Hace algunos años, Louis E. Brus, colaborador de Steigerwald, encontró que era posible mantener suspendidas en una solución partículas de seleniuro de cadmio, con diámetros comprendidos entre los 10 y los 100 angstroms. Más tarde, Brus y Steigerwald lograron recubrir las partículas con película de polímeros orgánicos que posteriormente se secaban y moldeaban en torno a los átomos. Estos recubrimientos impedían que los átomos se recombinaran cuando se extraían de la solución. Galen D. Stucky y su equipo de la Universidad de California en Santa Bárbara están intentando incorporar las partículas atómicas en cuestión en el seno de zeolitas y vidrios porosos, que, de este modo, actúan como si se tratara de jaulas. Otros autores trabajan con metales y utilizan láseres para fabricar partículas metálicas pequeñas partiendo de materiales de gran tamaño.

Aunque los químicos fabriquen partículas o puntos cuánticos cuyo tamaño sea muchísimo menor que el conseguido por los litógrafos, existe un problema. Como Steigerwald dice, en cualquier conjunto de partículas los tamaños pueden variar, por término medio, en un 10 por ciento. Esta variación es excesiva para los dispositivos electrónicos u ópticos, que exigen la utilización de partículas con tamaños idénticos. "No sé cómo resolver ese problema", admite Steigerwald. "Existen muchas ideas al respecto, pero el mundo necesitará otras fuentes de inspiración."

Kerry J. Vahala, del Instituto de Tecnología de California, acaba de formar equipo con un químico y un ingeniero para tratar de construir un sistema capaz de separar partículas mediante un haz de láser. El plan consiste en producir partículas de semiconductor en el seno de una corriente gaseosa y, a continuación, hacer incidir un haz de láser sobre dichas partículas. Aquellas partículas



LA LUZ PROCEDENTE DE UN LASER DE IONES KRYPTON (*izquierda*) excita un láser de pozo cuántico emisor de superficie (situado sobre la platina del microscopio), que emite un haz de rayos infrarrojos (*derecha*) en este ensayo acometido en los Laboratorios Sandia. (El haz de infrarojos se hizo visible provocando que esta radiación atravesara un material de conversión óptica.) Un monitor de televisión muestra el perfil transversal del haz procedente del láser emisor de superficie.

cuyos niveles energéticos entren en resonancia con el láser se separarán del conjunto restante y, posteriormente, se depositarán sobre un sustrato adecuado.

Otro grupo de Bellcore ha seguido un camino diferente, basado en la tensión y las deformaciones que se producen cuando un material depositado sobre otro crece de una forma determinada. Sometieron los materiales a cierta tensión con el fin de crear en ellos regiones en las que el intervalo entre bandas fuera pequeño y, de este modo, se produjeran puntos cuánticos.

En cualquier caso, la mayoría de los investigadores continúan, para este fin, con las técnicas litográficas, aunque introduciendo algunas modificaciones. Henryk Temkin, de AT&T, fabrica puntos cuánticos a partir de diferentes compuestos semiconductores, fosfuro de indio y arseniuro de galio e indio, que emiten luz con una frecuencia que puede transmitirse a través de fibras ópticas con poca atenuación. Por otra parte, Temkin reduce la posibilidad de contaminación del material y produce puntos de mejor calidad; realiza todas las fases del proceso de fabricación en el interior de un medio único de alto vacío.

Hace dos años, Reed y Randall, de la Texas Instruments, consiguieron fabricar un punto cuántico individualizado. Como dice Reed, dado que estaba aislado de cualquier otro dispositivo, el punto en cuestión "no resultaba útil, pero funcionaba". Los métodos para conectar un conjunto de puntos cuánticos en un circuito o sistema de mayor tamaño han planteado siempre muchos problemas. Meditando sobre esta cuestión, Steigerwald señala que "se podrían fabricar partículas eléctricamente activas conectándolas a un polímero conductor. No se trata de fantasía científica, sino de una posibilidad futura".

Randall continúa explorando las uniones eléctricas y reconoce que "mi mayor preocupación consiste en obtener buenas conexiones con los puntos cuánticos". Si alcanzaran el éxito, él y sus colaboradores intentarían utilizar estos puntos cuánticos como elementos diferenciados en sistemas informáticos. Dado que cada uno de estos puntos puede poseer varios estados energéticos diferentes, un solo punto podría reemplazar a varios transistores convencionales.

Los puntos cuánticos facultarían también al equipo de Texas Instruments para desarrollar el automatismo celular, un nuevo sistema de interconexión. En vez de conectar los puntos de transistor uno tras otro formando una serie, cada punto estaría unido a cuatro de sus vecinos más próximos. La unión de los puntos situados en los bordes de la red permitiría el control de las operaciones de los puntos situados en el interior. Este hecho convertiría al circuito en un sistema en paralelo general, en el que muchos transistores operarían simultánea, no secuencialmente.

Aunque una arquitectura como ésta no resultaría rentable con los transistores actuales, podría tener sentido si se utilizaran puntos cuánticos. "La reducción en el tamaño de los dispositivos proseguirá en el futuro y finalmente nos conducirá a los puntos cuánticos", dice Randall. "Cuando lleguemos a ese destino, necesitaremos disponer de sistemas adecuados para el acoplamiento de estos puntos."

La respuesta a la pregunta de "cuándo se llegará a ese destino" es incierta, y ya no nos queda mucho tiempo. Los fabricantes de redes de telecomunicación piden láseres más eficientes para poder enviar más datos a través de las líneas. Los fabricantes de microcircuitos lamentan que la capacidad de los métodos litográficos convencionales para empaquetar más y más dispositivos en una sola pastilla alcanzará un nivel de saturación a finales de nuestra década. Ha llegado el momento de la innovación.

Juegos de ordenador

Un compendio de matemaltratos llegados aquí de todo el mundo

A. K. Dewdney

"En el intervalo de 176 años, el Missisippi Inferior ha acortado por sí mismo la longitud de su curso en doscientas cuarenta y dos millas, lo que supone en promedio un poquito más de una milla y tercio al año. Por lo tanto, cualquier persona ecuánime que no sea ciega o idiota... podrá ver que el pasado noviembre hizo justo un millón de años que la longitud del río Missisippi Inferior pasaba de un millón trescientas mil millas."

—MARK TWAIN,
Life on the Missisippi

Desde el pasado mayo de 1990, en que traté por primera vez en esta sección los abusos cometidos so capa de las matemáticas, han sido centenares los lectores que me han proporcionado ejemplos de delitos matemáticos, amén de toda suerte de timos y errores de bulto. Los reos han sido sobre todo grandes compañías, instituciones y departamentos de la administración y del estado. Pero, ¿se trata de abusos intencionados o accidentales? Dejaré que sean los lectores quienes juzguen por sí mismos mientras van echando una ojeada a este minicatálogo de horrores matemáticos. Tal vez artículos como éste logren alzar efímeras barreras contra la marea de ignorancia que amenaza asfixiarnos —la corrupción de los porcentajes, la retorsión de la lógica, la presentación indebida (cuando no perversa) de datos numéricos, la distorsión de las tablas y gráficos—; sin embargo, a la larga, únicamente los sistemas educativos pueden ofrecernos alguna esperanza de salvación.

Corremos el riesgo de incurrir en abusos matemáticos por no saber, no poder, o no querer utilizar debidamente los números y relaciones con que nos tropezamos a diario. Incluso quienes no estén muy versados en álgebra o geometría pueden evitar el abuso matemático si afrontan este tipo de situaciones con un sano escepticismo. Un lector de Florida se muestra particularmente escéptico con los vendedores de automóviles, sobre todo cuando se empeñan en convencer al cliente de que les resulta ventajoso financiar el pago del vehículo en lugar de liquidarlo al contado. He aquí la divertida conversación —que he aderezado un poco— entre Juan, vendedor, y Paco, un presunto comprador.

"Pero, Paco, vamos a ver, ¿por qué quieres pagar de un golpe el millón que te cuesta el coche?", pregunta Juan.

"Para ahorrarme los intereses de la financiación, claro está", responde Paco. "Yo aceptaría el 11 por ciento de vuestra financiación si no fuera porque el banco me cobra un interés menor."

"Pero, ¿no te das cuenta de que aún así puedes ahorrar dinero financiándolo?", contraataca Juan. Se vuelve a su terminal de ordenador, teclea en él unas cifras y pronto le presenta a Pepe una pulcra hoja salida de la impresora, donde dice que con una cuota mensual de 32.739 pts. durante 36 meses se puede amortizar el millón de pesetas con un costo de 178.604. La hojita afirma además que depositando el millón en una cuenta a plazo al interés del 7,5 por ciento le proporcionaría 251.446 pts. "Como te decía, Paco, si dejas el millón en una cuenta a plazo en lugar de hacer efectivo todo el pago ahora, puedes ganar 72.842 pts., que es la diferencia entre los dos costos", declara Juan.

Un poco desconcertado, Paco inquiere: "¿Y qué me dices de las 32.739 que he de pagar todos los meses? Esa cantidad tendría que salir de mi cuenta bancaria..."

"Bueno, claro, estás pagando el coche", tercia Juan, "pero esa cantidad no tiene por qué salir de tu cuenta a plazo. Puede venir de otros capítulos; a lo mejor puedes detraerla de tu sueldo mensual. Francamente, Paco, de todos a quienes se lo he explicado, eres el único que no entiende el sistema".

Paco, un poco intimidado y algo confuso, decide prudentemente volver a casa y pensárselo despacio. Toma una hoja de papel y va anotando en dos columnas las previsiones de ambos planes. Bajo la columna "plan de Juan", al cabo de tres años tiene 1.251.446. En la de su propio plan, Paco paga el coche al contado e invierte al 7,5 por ciento la cuota mensual de 32.739 pts. Al cabo de los tres años tiene 1.317.135. Ahora es cuando Paco se percata de la añagaza de Juan. Juan evitó tener en cuenta los intereses que rendirían las cuotas mensuales si fueran invertidas en una cuenta bancaria de capitalización.

La forma de incompetencia numérica a la que llamo "ceguera al interés compuesto" es causa de que más de un ojo se nuble cada vez que se menciona algún plan financiero donde haya que atender a los intereses. Varios lectores me comentaron diversos planes de jubilación de muy atractivo aspecto hasta que se tiene en cuenta la inflación, y no digamos ya, los impuestos. A decir verdad, estoy dispuesto a enunciar una ley de inflación e impuestos, que es preciso tomar medio en serio, pero sólo medio en broma:

Ley de retornos nulos
Retorno de inversión = pérdida por inflación + impuestos

Supongamos, por ejemplo, que la tasa de interés sea del 9 por ciento, que la inflación sea del 5,5 por ciento y que las cargas fiscales sobre las ren-



tas del capital asciendan al 35 por ciento. Si uno pone 1.000.000 de pts. a un plazo de un año, recibirá 1.090.000. Pero el mordisco del 35 por ciento con que los impuestos gravan las 90.000 pts.de interés nos dejan solamente 58.500. ¡Y ahora, a causa de la inflación, las 1.058.500 pts. restantes equivaldrían solamente a 1.000.283!

Mas no basta comprender el interés compuesto para quedar vacunado contra abusos. Los publicitarios, y no digamos los vendedores, se distinguen por su habilidad para manipular los números y escamotear la información que pudiera restar atractivo a sus propuestas. Un anuncio televisivo proclama la marca de analgésico preferida por una muestra aleatoria de médicos detenidos a su pesar en una isla desierta. Según el patrocinador de la marca X, más del 50 por ciento de los infortunados galenos la prefieren a la marca Y, segunda en popularidad. Lo que el anuncio no revela explícitamente es que en la encuesta se les ofreció a los médicos una tercera opción, Z, que no era más que una variante de Y. Como dice mi comunicante, la preferencia por el elemento activo de las marcas Y y Z superaba a la de la marca X aproximadamente en la misma proporción en que X era preferida a Y.

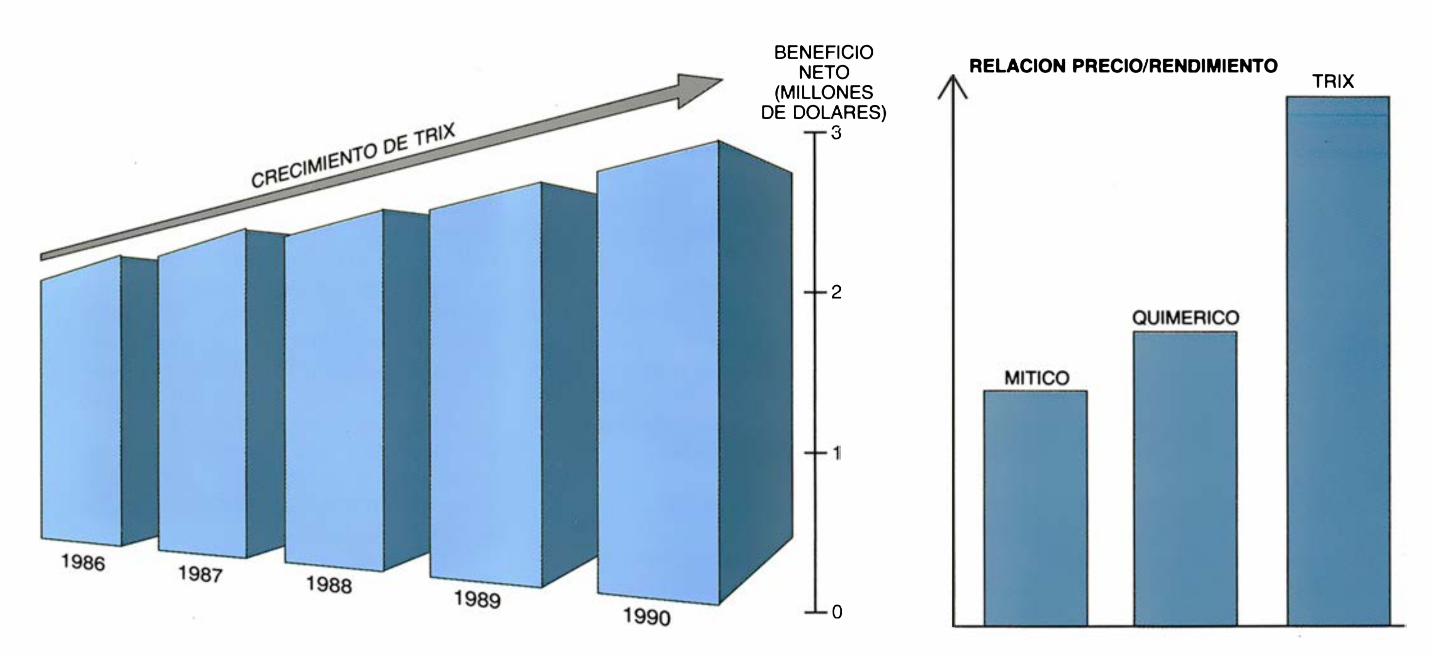
Los anunciantes pueden demostrar prácticamente lo que se quiera a base de ocultar información. Hace algunos años, el anunciante de una conocida marca de whisky, A, proclamaba que en una prueba "a ciegas" la mitad de los habituales de otra marca prefería en realidad la marca A. "¿Qué significa esto en realidad?", pregunta un corresponsal de Estocolmo. "El resultado viene a ser el mismo que si la elección se hiciese lanzando una moneda a cara o cruz", me dice. Lo cual no significa que el resultado de la prueba sea aleatorio, pero sí que debería despertar sospechas. Después de todo, son muchos los consumidores de bebidas que son realmente incapaces de distinguir su marca favorita al degustarla junto a otras marcas competidoras, por muy convencidos que estén de lo contrario. El año pasado, la Canadian Broadcasting Company pasó un documental sobre bebedores de cerveza que se tenían por *connoiseurs* de las diversas marcas. En pruebas a ciegas, los *cognoscenti* confundían sistemáticamente una marca con otra.

Algunos fabricantes de bebidas sacan partido de la insensibilidad de las papilas gustatorias del público, sobre todo en refrescos y bebidas frías. Por ejemplo, un conocido fabricante de bebidas no alcohólicas desafía a los consumidores a decidir cuál de dos marcas prefieren en una prueba de degustación, si la suya o la de la competencia. Dada la semejanza de sabor de una y otra, alrededor de la mitad de bebedores de la marca rival dicen preferir el sabor de la anunciada. La publicidad de esta clase es eficaz: la mayoría de los degustadores ignoran el fenómeno de confusión de sabores. A decir verdad, el hecho mismo de tener una marca "favorita" es casi segura garantía de que no se está dispuesto a admitir que tal preferencia se funda en cualquier cosa menos en el sabor.

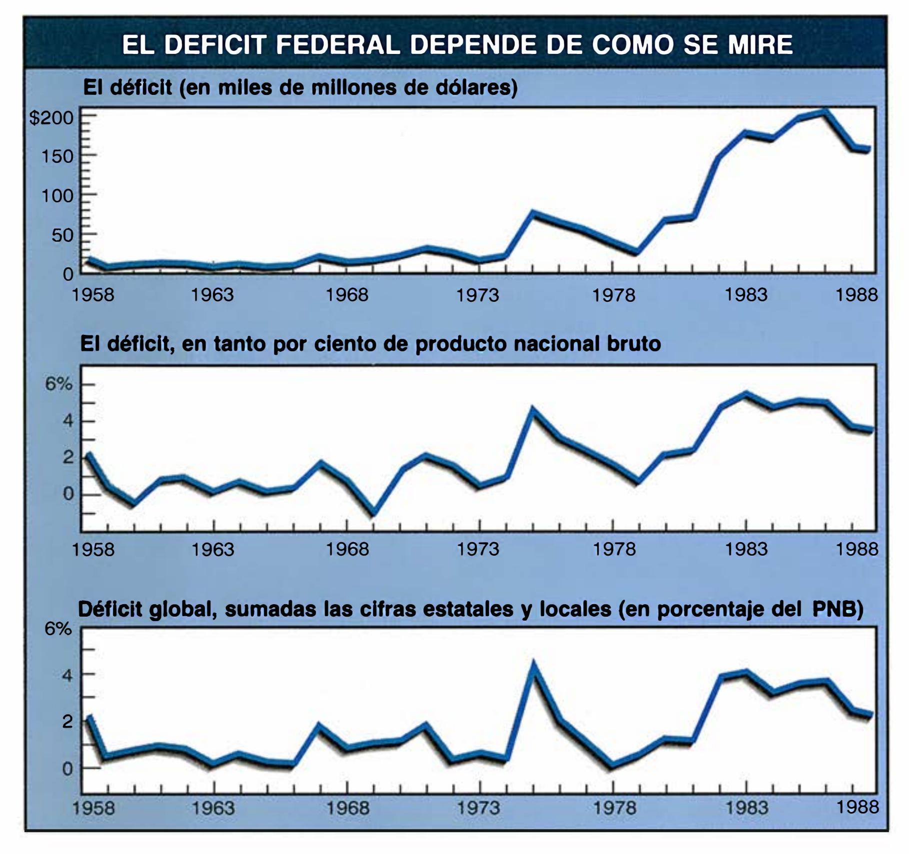
El público podría sentirse inclinado a creer que son las grandes sociedades anónimas y las agencias de publicidad quienes cometen los mayores abusos matemáticos; pero, a juzgar por el correo que he recibido, los organismos oficiales y los medios de comunicación no les andan a la zaga. Un anuncio publicado en un periódico sueco clama, "¡SEGURIDAD SEIS VECES MAYOR! El año pasado se ahogaron 35 personas por accidentes de navegación en pequeñas embarcaciones. Solamente cinco de ellas llevaban chaleco salvavidas; las demás no. Así pues, cuando navegue lleve siempre puesto el chaleco salvavidas."

Nadie discute la laudable intención del mensaje, pero ¿qué podemos verdaderamente deducir de estas cifras?, se pregunta un lector danés. ¿Es realmente posible que haya seis veces más suecos navegando a vela sin chaleco salvavidas que con chaleco puesto? Resulta irónico pensar que si así fuera y si la proporción de suecos que los utilizasen aumentase grandemente, seguramente se invertirían las proporciones en las estadísticas del ejemplo. En este caso el abuso podría denominarse "ayuno estadístico": sencillamente, no se nos proporciona información bastante para que los números que se nos facilitan resulten convincentes. Y en lugar de dárnosla, los medios de comunicación se limitan a repetir alegremente lo que las diversas agencias y organismos les quieren contar.

En cuanto uno reúne todas las pie-



1. Trixsa: crecimiento rápido que oscurece a la competencia.



2. Acertijo: ¿En qué años tuvieron los EE.UU. un PNB negativo?

zas, resulta fácil observar cómo pueden los medios de comunicación aderezar los números e incurrir en una forma de maltrato demasiado frecuente, al que llamo inflación numérica. Desde California, un lector me ha hecho saber documentadamente de declaraciones sobre las diversas causas de los accidentes mortales de circulación. Hélas aquí: cocaína, 20 por ciento (según un diario neoyorquino); marihuana, 25 por ciento (DEA, Agencia antidroga estadounidense); alcohol, 50 por ciento (Servicio de vigilancia de carreteras de California); somnolencia, 35 por ciento (investigadores del sueño); exceso de velocidad, 85 por ciento (Comité nacional de seguridad de transportes); fumar, 50 por ciento (Comité nacional de seguridad del tráfico por carretera); suicidio, 35 por ciento (investigadores del suicidio); fallos mecánicos, 20 por ciento (Departamento de vehículos motorizados del estado de Nueva York).

Admito de buen grado que las categorías distan de ser mutuamente excluyentes, por lo que no debe sorprender que los porcentajes sumen más de 100. Además, algunos de los porcentajes pueden ser más aplicables a las carreteras que a las ciudades. No obstante, de dar crédito a estos porcentajes, no resultaría nada sorprendente que cada año hubiera varios millares de conductores, que en sus esfuerzos por darse muerte, esnifan unas rayas, se fuman varios porros, se vierten gaznate abajo medio litro de licor, se fuman una cajetilla y por fin, a bordo de una vieja perola de hace 20 años, aceleran hasta rebasar las limitaciones de velocidad y caer dormidos sobre el volante.

A mucha gente le cuesta prestar atención a los números grandes, sobre todo si además intervienen porcentajes. Más de uno habrá guiñado los ojos al leer el titular que me envió un lector de Chicago: "SIETE DE CADA DIEZ ITALIANOS HAN COMETIDO ADULTERIO". Después de todo, el 49 por ciento de los hombres y el 21 por ciento de las mujeres que contestaron a una encuesta confesaron haber tenido relaciones extramaritales. Y como todo el mundo sabe,

$$49 + 21 = 70.$$

Pero al sumar directamente los porcentajes estamos despreciando las proporciones relativas de hombres y mujeres en la población italiana (suponiendo, lo que no es poco, que la muestra encuestada fuese representativa). Si tal proporción fuese más o menos de mitad y mitad, entonces el 24,5 por ciento de la población italiana estaría formada por hombres que han sido infieles a sus dulces esposas y el 10,5 por mujeres que lo han sido a sus maridos. Estos porcentajes sí se pueden sumar ahora, porque los dos grupos representan partes disjuntas de la misma población:

$$24{,}5 + 10{,}5 = 35.$$

En este caso es probable que el abuso matemático haya sido en gran medida inconsciente, aunque algo pudo pesar cierta tendencia a la exageración típica de los medios de comunicación. Es probable que los infractores se encontrasen bajo los efectos de la misma forma de incompetencia numérica que las víctimas: la ratioitis. Pues el "anaritmetismo" se da lo mismo en porcentajes que en fracciones, proporciones y promedios. Ninguno de estos valores puede sumarse o multiplicarse con el mismo indolente abandono que los números ordinarios. Es enfermedad que ataca incluso a gentes instruidas. Y a los educadores. Un corresponsal de Saratoga cita un informe sobre resultados académicos: "Aunque en la década de los setenta las puntuaciones de los exámenes descendieron casi un 60 por ciento, desde entonces hasta hoy se han remontado más de un 70 por ciento".

Lo mismo que antes, la tendencia es la de sumar las dos cifras, como si de números ordinarios se tratase: restar el 60 por ciento y sumar el 70. Pero apliquemos ahora la caída y subida proclamadas a un valor concreto, 85 pongamos por caso. Si el número 85 pierde el 60 por ciento de su valor decrece en 51, y se convierte en 34. Si ahora el número 34 se "remonta" un 70 por ciento ganará un 70 por ciento de su valor, o sea, 23,8 unidades, y sube hasta 57,8, muy por debajo del 85 primitivo. Mi comunicante llama "pernicioso" a este abuso de los porcentajes y escribe: "Sin duda debería disculpársenos que no alcancemos a comprender que las puntuaciones sean ahora menos del 70 por ciento de lo que inicialmente fueron". ¡Es lastimoso que quienes informan sobre los progresos en educación sean, ellos mismos, botón de muestra de sus deficiencias.

Quisiera dejar a los lectores dos tipos de divertimentos, pequeños y sencillos ejercicios con los que tonificar esos músculos mentales concebidos para tensarse y entrar en acción

en la jungla de la vida y alertarnos de la falsedad o la estupidez. El primer tipo de divertimento se ocupa del abuso de diagramas, que es la distorsión y manipulación de los datos, sea deliberada o intencional, por medios gráficos. ¿Cuántos lectores podrán detectar el abuso cometido en cada uno de los diagramas que ilustran este artículo? A primera vista, cada uno de los diagramas puede ofrecer aspecto inocente, pero una segunda mirada debería poner en acción nuestro espíritu crítico.

La segunda clase de divertimento procede de Robert M. Martin, que enseña un curso sobre la lógica de la probabilidad y sobre razonamiento científico en la Universidad Dalhousie, en Halifax. Martin ha zancadilleado más de una vez a sus alumnos —estudiantes universitarios— con preguntitas ingenuas, como las tres siguientes:

- Se toman de la baraja un as, un rey y una reina, se mezclan y se dan cara abajo sobre la mesa. Alguien, por accidente, hace caer al suelo una de ellas, completamente al azar. La carta caída queda cara abajo. ¿Cuál es la probabilidad de que al tomar una carta de la mesa sea un as?

- Mucha gente trabaja ocho horas diarias, o sea, la tercera parte de un día de 24 horas, por lo que estas personas trabajan en realidad alrededor de un tercio de los 365 días del año, casi 122 días. Pero como no se trabaja los fines de semana, o sea, dos días por semana, hay que descontar 104 días al año. Al restar los 104 días de los 122 nos quedan solamente 18 días. Dado que además, hay que descontar 30 días de vacaciones al año, ¡nadie trabaja!

- Una botella de licor y la caja de presentación cuestan 1100 pesetas. Si la botella cuesta 1000 pesetas más que la caja, ¿cuánto cuesta la caja? Al responder (mal) esta pregunta, un estudiante escribió zumbón, "¡Qué difícil!"

Los ejemplos de abusos anteriores no son sino meros síntomas de una generalizada ignorancia de las matemáticas y, aún más, de la ciencia en su conjunto. De entre las naciones industrializadas, parece como si la enfermedad hubiera alcanzado máxima virulencia en los Estados Unidos, país donde los escolares obtienen por término medio puntuaciones más bajas que sus compañeros de otros países en pruebas de examen normalizadas. Tenemos plena conciencia de la crisis de los sistemas educativos pero, ¿qué podemos hacer?

Lynn Arthur Steen es matemático del St. Olaf College, de Northfield, Minnesota, presidente de la Sociedad Matemática Americana y renombrada autoridad en educación matemática desde el jardín de infancia hasta la docencia universitaria. Según Steen, los niños salen de la escuela primaria e incluso de la secundaria superior en estado de confusión matemática; o, en el mejor de los casos, convertidos en meros manipuladores de símbolos, en autómatas que aplican mecánicamente las fórmulas que han aprendido. Son muy pocos los estudiantes estadounidenses que poseen lo que Steen llama "sentido común numérico" o "buen sentido simbólico".

El "buen sentido" numérico es reflejo justo de la clase de competencia que se precisa para combatir el abuso y maltrato de las matemáticas. Comporta la capacidad para comprender magnitudes relativas, el crecimiento en progresión geométrica —el llamado "crecimiento exponencial"— y el significado de razones y porcentajes, por sólo mencionar unas pocas destrezas. Los estudiantes poseerán buen sentido simbólico si logran reconocer que un resultado matemático es incorrecto porque no encaja con el valor estimado de antemano. Tales alumnos son capaces de pensar algorítmicamente y de abarcar con relativa facilidad procesos abstractos sencillos.

En Estados Unidos, el Congreso Nacional de Profesores de Matemáticas (NCTM) ha recomendado niveles mínimos y normas de evaluación para los escolares, y la Comisión de Educación de Ciencias Matemáticas, dependiente de la Academia Nacional de Ciencias, ha estado promoviendo activamente la educación matemática en los Estados Unidos. Aunque es cierto que unas cuantas escuelas están haciendo un magnífico trabajo en la enseñanza de las matemáticas, señala Steen, la mayoría están hundiéndose en un agujero negro creado por la indiferencia pública. Si los Estados Unidos pudieran apasionarse por la crisis en la educación matemática con tanto fervor como ha entendido los problemas ambientales, prosigue Steen, quizás pudiéramos dar la vuelta a la situación.

BIBLIOGRAFIA COMPLEMENTARIA

EVERYBODY COUNTS: A REPORT TO THE NATION ON THE FUTURE OF MATHEMATICS EDUCATION. National Academy Press, 1989.

Juegos matemáticos

Con ocasión de una visita a Euler, el lector descubrirá el interés de los números perfectos

Ian Stewart

Es posible que el lector haya oído hablar del torsionador de espacio-tiempo que tengo en el jardín, tras las varas de frambuesas. No suelo hablar de él, no sea que se inquieten los vecinos, pero de cuando en cuando me viene muy bien. Yo lo utilizo sobre todo para viajar al planeta Ombilicus, que parece ser su destino "por defecto", como dicen los informáticos; pero he descubierto que a costa de una intensa concentración puedo dirigirme prácticamente a cualquier lugar y tiempo que desee. Lo malo es que a veces, como me distraiga, acabo en épocas y lugares completamente distintos de los que había pensado.

Era evidente que había ido a parar a la Rusia zarista. La cuestión era: ¿por qué había acabado en la Rusia zarista? ¿Y quién era aquel ciego sentado cerca del fuego?



1. Leonhard Euler, nacido en Basilea en 1707, fallecido en San Petersburgo en 1783. Se le reconoce el mayor número de descubrimientos matemáticos originales, por encima de cualquier otro, antes o después que él.

Me esforcé por reconstruir mis pensamientos. Me había encaminado hacia el Laboratorio de Infinormática de Ombilicus con el fin de convencer al Dr. Bunnydew de que "corriera" para mí un programa infinito de ordenador. (Es favor que me hace de vez en cuando, pero no puedo publicar ninguno de los resultados que obtengo porque nadie más podría verificarlos... a menos que dispusiera de un torsionador de espacio-tiempo tras las frambuesas, y en tal caso también ellos tendrían que guardar el secreto.) El programa en cuestión se proponía zanjar de una vez por todas el problema de Euler concerniente a la existencia o inexistencia de números perfectos impares.

Ya estaba claro. Era Euler.

En 1727, Leonhard Euler, el más prolífico matemático de todos los tiempos, había sido llamado a la Academia de San Petersburgo. La Academia había sido fundada por Yekaterina Alekseyevna, también conocida por Catalina I de Rusia, la cual deseaba que Euler fuese matemático permanente de su Academia. Catalina I reinó desde 1725 hasta 1727 y murió el mismo día en que Euler ponía el pie en Rusia, con lo que éste se vio obligado a pasar políticamente inadvertido durante largos años. Es más que posible que salvase el cuello hundiendo la cabeza en las matemáticas; lo que Euler perdió con ello ha redundado en ganancia de la comunidad matemática. Euler permaneció en Rusia hasta 1742; se trasladó entonces a Berlín y finalmente regresó a San Petersburgo (hoy Leningrado) en 1766 durante el reinado de Catalina la Grande. Allí murió en 1783. Y se quedó ciego hacia 1770 (aunque ello no influyó en su producción matemática, que seguía viendo con su "ojo mental"). Por lo tanto, yo me encontraba en San Petersburgo, en algún momento del reinado de Catalina la Grande.

Euler dijo unas palabras en ruso. Al no responder yo, volvió a intentarlo en alemán. "¿Quién está ahí?" Los ciegos tienen el oído muy fino; sin duda me estaba oyendo respirar. En mi alemán, no muy fluido, procuré explicarle que yo era un matemático procedente del futuro, que había aterrizado accidentalmente en San Petersburgo yendo de camino a una estrella distante. Sin duda me tomó por un chiflado, aunque siguió siendo cortés, quizá por seguirme la corriente. En un esfuerzo por acreditar mi personalidad le referí diversos descubrimientos matemáticos recientes, como el contraejemplo descubierto por Noam Elkies a su conjetura de que la suma de tres cuartas potencias no puede ser una cuarta potencia. Mi explicación le mantuvo interesado, pero creo que lo que verdaderamente le convenció fue el tacto para él desconocido que tenían mis ropas del siglo XX. Después de eso pareció fascinado por una nueva versión de la mecánica que podría permitir los viajes por el tiempo. Euler hallaba inspiración en cualquier cosa.

Para distraerle de tan peligrosa dirección —supongamos que tuviera éxito; ¿qué sucedería entonces con la historia venidera?— le mencioné que en Ombilicus me proponía verificar su conjetura de que no puede existir un número perfecto que sea impar. Lo cual dio pie a una conversación apasionante, pues Euler había trabajado en el problema, aunque no había logrado resolverlo. No obstante, sí había echado los cimientos para casi todo el trabajo posterior.

El lector sabrá seguramente qué son los números perfectos. Son números iguales a la suma de sus divisores propios (el término clásico era "de sus partes alícuotas"). Los dos primeros números perfectos son

$$6 = 1 + 2 + 3, \text{ y}$$
$$28 = 1 + 2 + 4 + 7 + 14$$

Resulta matemáticamente más conveniente considerar que cada número N es divisor de sí mismo y definir $\sigma(N)$ como la suma de todos los divisores de N, incluido el propio N. Los números perfectos son entonces los que verifican la condición $\sigma(N) = 2N$. (El símbolo σ es la letra griega 'sigma', que representa sumación.) La razón de también contar a N entre los divisores de N en la definición de $\sigma(N)$ es que de esta forma $\sigma(N)$ se torna *multiplicativa* en el siguiente sentido: si M y N son números primos entre sí (o como también se dice, *coprimos*), esto es, sin divisores comunes mayores que 1, entonces $\sigma(NM) = \sigma(N)\sigma(M)$. El recuadro A explica la razón de que se verifique esta ecuación. La relación es *falsa* si los nú-

meros poseen algún divisor común: por ejemplo

$$\sigma(2) = 1 + 2 = 3$$
$$\sigma(4) = 1 + 2 + 4 = 7$$
$$\sigma(8) = 1 + 2 + 4 + 8 = 15$$

pero $15 \neq 3 \times 7$, con lo cual $\sigma(8) \neq \sigma(2)\sigma(4)$.

El primer teorema general atinente a los números perfectos fue demostrado por Euclides: si N es de la forma $2^{p-1}(2^p - 1)$, donde 2^{p-1} es primo, entonces N es perfecto. La multiplicatividad de σ permite una demostración sencilla. Dado que $2^p - 1$ es primo, sus únicos divisores son 1 y $2^p - 1$, así que

$$\sigma(2^p - 1) = 1 + (2^p - 1) = 2^p$$

Por otra parte,

$$\sigma(2^{p-1}) = 1 + 2 + 4 + \ldots + 2^{p-1} = 2^p - 1.$$

Así, pues,

$$\sigma(N) = \sigma[2^{p-1}(2^p - 1)] = \sigma(2^{p-1})\,\sigma(2^p - 1) = (2^p - 1)2^p = 2\,[2^{p-1}(2^p - 1)] = 2N.$$

Hay en este caso un "intercambio de primos" muy elegante entre 2 y $2^p - 1$: el divisor 2^{p-1} de N aporta un factor 2^p a $\sigma(N)$, mientras que el factor $2^p - 1$ de N aporta un factor 2^p a $\sigma(N)$.

Por ejemplo, si $p = 2$ entonces $2^2 - 1 = 3$ es primo, y el resultado de Euclides muestra que $2 \times 3 = 6$ es perfecto. Si $p = 3$ entonces $2^3 - 1 = 7$ es primo y $2^2 \times 7 = 28$ es perfecto. Así, el teorema recupera los dos ejemplos que hemos visto ya. Sin embargo, cuando $p = 4$, $2^4 - 1 = 15$ *no* es primo sino compuesto, pues es producto de 3 por 5. En este caso la fórmula no genera un número perfecto. Tenemos

$$2^{p-1}(2^p - 1) = 2^3(2^4 - 1) = 8 \times 15 = 120,$$

pero por la multiplicatividad

$$\sigma(120) = \sigma(8)\sigma(3)\sigma(5) = 15 \times 4 \times 6 = 360.$$

Así pues, $\sigma(120) = 3 \times 120$, no 2×120. Decimos entonces que 120 es *triperfecto*.

Los números $M_p = 2^p - 1$ se llaman *números de Mersenne* y en el caso de ser primos, *primos de Mersenne*. Se conocen en la actualidad 30 primos de Mersenne, a menos que alguien haya descubierto uno nuevo mientras yo estaba de espaldas, lo cual no es inverosímil, porque cada vez que se construye un ordenador más rápido normalmente se le afilan los dientes hallando un nuevo primo de Mersenne. No es difícil demostrar que si M_p es primo, entonces p es primo. No obstante, M_p puede ser compuesto aun cuando p sea primo: el primer ejemplo es $2^{11} - 1 = 2047 = 23 \times 89$.

Euler demostró un recíproco del teorema de Euclides: si N es un número perfecto par, entonces $N = 2^{p-1}(2^p - 1)$ con $2^p - 1$ primo. Si dejamos de lado el problema de decidir cuándo es primo un número de Mer-

Primos de Mersenne

Los números $M_p = 2^p - 1$ son primos para los siguientes valores de p:

2	61	2282	21701
3	89	3217	23209
5	107	4253	44497
7	127	4423	86243
13	521	9689	132049
13	607	9941	216091
19	1279	11213	
31	2203	19937	

A: Multiplicatividad de la función "suma de divisores"

A modo de ejemplo, examinemos el caso $120 = 8 \times 15$. Los divisores de $8 = 2^3$ son 1, 2, $2^2 = 4$, y $2^3 = 8$; los hemos escrito a lo largo de la fila de encabezamiento de la tabla. Los divisores de 15 son 1, 3, 5, y $3 \times 5 = 15$; escribámoslos en columna, a la izquierda. Al desarrollar el producto $(1 + 2 + 2^2 + 2^3)(1 + 3 + 5 + 3 \times 5)$, obtenemos la suma de todos los términos de la tabla; pero estos términos son precisamente los divisores de 120, pues cada uno de sus sumandos corresponde a la elección de ciertos divisores de $120 = 2^3 \times 3 \times 5$ y a la agrupación de los mismos en cierta potencia de 2 multiplicada por algo formado a partir de 3 y 5. Así pues, $\sigma(120) = \sigma(8)\sigma(15)$.

Divisores de 8 → / **Divisores de 15** ↓	1	2	2^2	2^3
1	1	2	2^2	2^3
3	3	2×3	$2^2 \times 3$	$2^3 \times 3$
5	3	2×5	$2^2 \times 5$	$2^3 \times 5$
3×5	3×5	$2 \times 3 \times 5$	$2^2 \times 3 \times 5$	$2^3 \times 3 \times 5$

divisores de 120

La demostración subsiste esencialmente idéntica para cualquier par de enteros m y n que sean primos entre sí. Sin embargo, cuando m y n tienen algún divisor común el método anterior falla. Por ejemplo, si sustituimos 15 por $14 = 2 \times 7$, la tabla se convierte en:

Divisores de 8 → / **Divisores de 14** ↓	1	2	2^2	2^3
1	1	2	2^2	2^3
2	2	2^2	2^3	2^4
7	7	2×7	$2^2 \times 7$	$2^3 \times 7$
2×7	2×7	$2^2 \times 7$	$2^3 \times 7$	$2^4 \times 7$

divisores de 112

pero algunos de los divisores de 112 están repetidos: cada uno de los divisores 2, 2^2, 2^3, 2×7, $2^2 \times 7$, $2^3 \times 7$ aparece dos veces. Así pues, mediante esta técnica lo más que podemos demostrar es que si m y n tienen un divisor es que $\sigma(mn) < \sigma(m)\sigma(n)$.

Una consecuencia directa de la multiplicatividad es que si $l, m, \ldots, n$ son primos dos a dos, entonces $\sigma(lm\ldots n) = \sigma(l)\sigma(m)\ldots\sigma(n)$.

¿Se podría utilizar el resultado anterior para deducir una fórmula que exprese $\sigma(N)$ a partir de la descomposición en factores primos $N = p^a q^b \ldots r^c$?

senne, el resultado anterior caracteriza del todo a los números perfectos pares. Euler demostró también que un número perfecto impar, de existir alguno, tiene que ser de la forma $N = p^{4k+1}s^2$, donde p es un número primo de la forma $4m + 1$ y s no es divisible por p. Las demostraciones de los resultados de Euler, que son nuevamente consecuencia de la multiplicatividad de σ, se ofrecen en el recuadro B. Ciertos acontecimientos clave de la historia de los números perfectos, entre los que figuran algunos de los descritos detalladamente debajo, se muestran en el recuadro C.

"¿Ha descubierto alguien un número perfecto *impar*, o demostrado su inexistencia?", me preguntó Euler.

"No", le respondí. "La cuestión es delicada."

"Tal me lo pareció", repuso con sequedad. "Muchas horas he consagrado a esta cuestión, pero con parcos frutos."

"Confío en que no gastara demasiadas tanteando ejemplos", dije yo. "En el futuro del que vengo, dos australianos —un científico informático llamado Richard Brent y un matemático de nombre Graeme Cohen— han demostrado hace poco que un posible número perfecto impar habría de tener al menos 161 cifras. Y además han anunciado que la cota inferior puede elevarse a 301 dígitos."

"¿Qué significa australianos?", preguntó Euler con perplejidad en la voz. "¿Quiere vuesa merced decir 'austríacos'? Tengo entendido que Viena es un centro de actividad matemática..."

"Australia es una tierra nueva, situada muy al sur, aún no explorada en su tiempo."

"Y un científico informático, ¿qué es?"

"Es una persona que se vale de máquinas muy complejas para efectuar con rapidez cálculos complicados, o bien es una persona que diseña y proyecta tales máquinas, o que piensa acerca de la teoría de tales máquinas."

"Me resulta imposible creer que máquina alguna, por muy potente que sea, pueda examinar 10^{300} casos individuales", atajó Euler. "Yo doy por supuesto que las ideas matemáticas eran utilizadas —bueno, disculpad, *serán*— para reducir el número de casos a volúmenes tratables."

"En efecto, así es", dije. "La verdad es que lo que más interesa es el proceso matemático utilizado y no el resultado. Si hubieran llegado a encontrar un número perfecto impar... bueno, eso hubiera sido algo muy notable. Pero mientras que resulta grato saber que no existe ninguno menor que 10^{300}, resulta mucho más estimulante ver cómo se podría llegar a demostrar una cosa así."

"Entonces, quizá debiera vuesa merced dejarse de circunloquios y proceder a explicar el método", me reprendió.

Le rogué que perdonase mis digresiones. "El método resulta más sencillo de comprender mediante ejemplos", dije. "Voy a mostraros dos casos: una demostración de que no existe ningún número perfecto impar menor que 10^4, que permite establecer los principios básicos pero es demasiado corta para ser plenamente representativa, y después, una demostración de que no existe ningún número perfecto impar menor que 10^6."

Todo se funda en dos ideas. Una es la multiplicatividad de la función "suma de divisores"; la otra es la forma euleriana estándar $N = p^{4k+1}s^2$ para los números N que sean perfectos e impares. Expresemos s como producto de distintas potencias de números primos $q^a r^b\ldots$, de suerte que $s^2 = q^{2a}r^{2b}\ldots$ y tendremos $N = p^{4k+1}q^{2a}r^{2b}\ldots$ Llamemos *factores* a los números p, q, r y sucesivos; llamemos *componentes* a sus potencias, p^{4k+1}, q^{2a}, r^{2b} etcétera, y denominemos *exponentes* a $4k + 1$, $2a$, $2b$... Daré a p el nombre de primo especial porque es el único que aparece elevado a exponente impar.

Supongamos que N sea impar y fijémonos en la ecuación $\sigma(N) = 2N$ que determina a los números perfectos. Por la multiplicatividad,

$$\sigma(N) = \sigma(p^{4k+1})\sigma(q^{2a})\sigma(r^{2b})\ldots$$

con lo cual

$$2N = \sigma(p^{4k+1})\sigma(q^{2a})\sigma(r^{2b})\ldots$$

Se siguen de aquí dos consecuencias inmediatas:

– Cada divisor impar de $\sigma(N)$ tiene que ser también un divisor de N, y recíprocamente.

– Exactamente uno de los números $\sigma(p^{4k+1})$, correspondiente al primo especial, o $\sigma(q^{2a})$, correspondiente a una de las componentes, q^{2a}, es divisible por 2, pero no por ninguna de las sucesivas potencias de 2.

La idea consiste en servirse de estas observaciones para deducir, a partir de la presencia de ciertos divisores de N, la presencia de otros divisores. El proceso continúa hasta que el número de divisores es tan grande que N rebasa la cota de 10^4. El procedimiento se llama "método de la cadena de divisores" porque va construyendo más y más divisores de N a partir de cada hipotética componente.

"Sí, ya conozco este método", dijo

B: Dos demostraciones eulerianas de perfección

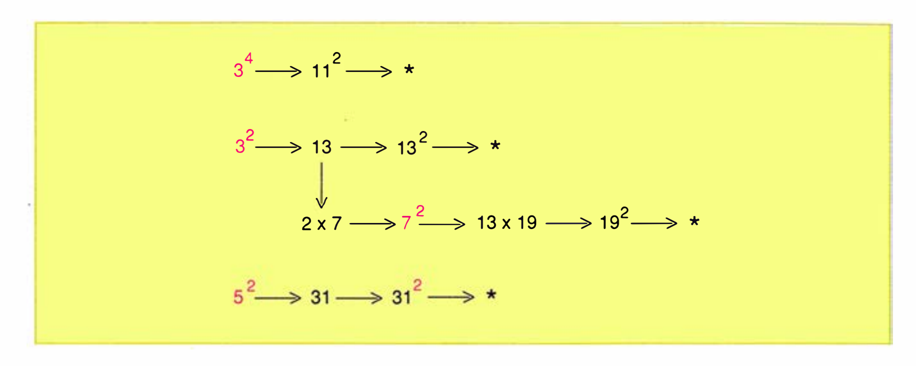
1 Los números perfectos pares tienen que ser de la forma $2^{p-1}2^p - 1$, siendo 2^p-1 un número de Mersenne.

Sea N un número perfecto par, que expresaremos como $N = 2^n b$, siendo $n \geqslant 1$ y b impar. Entonces $\sigma(N) = \sigma(2^n)\sigma(b) = (2^{n+1} - 1)\sigma(b)$. Dado que N es perfecto, $\sigma(N) = 2N = 2^{n+1}b$. Así pues, $b/\sigma(b) = (2^{n+1} - 1)/2^{n+1}$. La fracción del segundo miembro es irreducible, por lo que existe un c tal que $b = (2^{n+1} - 1)c$, $\sigma(b) = 2^{n+1}c$. Si $c > 1$ entonces b posee al menos los divisores 1, b, c, con lo cual $\sigma(b) \geqslant 1 + b + c = 1 + 2^{n+1}c + 1 > 2^{n+1} > \sigma(b)$, que es absurdo. Por consiguiente, $c = 1$ y $N = 2^n(2^{n+1} - 1)$. Finalmente, si $2^{n+1} - 1$ no es primo, tendrá divisores distintos de sí mismo y de 1, de donde $\sigma(2^{n+1} - 1) > 2^{n+1}$, que también es absurdo. Pongamos ahora $n + 1 = p$.

2 Los números perfectos impares tienen que ser de la forma $p^{4k+1}s^2$, siendo p un primo de la forma $4m + 1$ y siendo s y p números primos entre sí.

Escribamos $N = p^a q^b \ldots r^c$, siendo $p, q, \ldots, r$ números primos y distintos. Entonces $\sigma(N) = \sigma(p^a)\sigma(q^b)\ldots\sigma(r^c)$ y también igual a $2N$. Exactamente uno de los números $\sigma(p^a)$, $\sigma(q^b), \ldots, \sigma(r^c)$ es doble de un número impar y los demás han de ser impares. Elijamos la notación de manera que $\sigma(p^a)$ sea doble de un número impar. Ahora bien, $\sigma(q^b) = 1 + q + q^2\ldots + q^b$ es impar cuando b es par y par cuando b es impar; por lo tanto b es par. El mismo razonamiento vale para todos los factores siguientes; hemos demostrado pues que $N = p^a s^2$ para $s = q^{b/2}\ldots r^{c/2}$. Análogamente, a es impar. Si $a = 4k - 1$, entonces $\sigma(p^a) = 1 + p + \ldots + p^a$ consta de $4k$ sumandos. Las $2k$ potencias pares presentan la forma $4l + 1$ (para l diferentes) y su suma posee la forma $4m + 2k$; las $2k$ potencias impares son o bien todos de la forma $4l + 1$, y entonces su suma es $4n + 2k$, o bien son todos de la forma $4l - 1$ y su suma es $4n - 2k$; en todos los casos, $\sigma(p^a)$ es divisible por 4, lo que es imposible, por ser el doble de un número impar. Por consiguiente a es de la forma $4k + 1$.

Finalmente, si p es de la forma $4m - 1$ entonces $\sigma(p^{4p+1})$ es una suma de $4k + 2$ términos que van dejando alternativamente restos +1 o –1 al dividirlos por 4, así que ella misma es divisible por 4; contradicción. Luego p es de la forma $4m + 1$.



2. Arboles de divisores para demostrar la inexistencia de números perfectos impares menores que 10^4. Las cifras en color corresponden a posibles componentes. Las flechas indican deducciones de nuevos factores de N *a partir de los presupuestos; las flechas tras el 13 y el 61 parten de los hipotéticos primos especiales. Los asteriscos (*) indican contradicciones, por ser demasiado grande la componente hallada.*

Euler. "Por ejemplo, si el primo especial fuese 13, con exponente 1, entonces $\sigma(13)$ dividiría a $2N$. Pero $\sigma(13) = 1 + 13 = 14 = 2 \times 7$, por lo que 7 tiene que dividir a N. Puesto que 7 no es el primo especial, alguna potencia par, como 7^2 o 7^4 es una componente de N. Si fuese 7^2, por ejemplo, entonces $\sigma(7^2) = 1 + 7 + 49 = 57$ dividiría a N. Pero $57 = 3 \times 19$, y así, por el mismo argumento, 3^2 u otra potencia mayor de 3, y 19^2 u otra potencia superior de 19 dividiría a N..., etcétera. En este caso tenemos ya que N es múltiplo de $13 \times 7^2 \times 3^2 \times 19^2 = 2.069.613$, que es mayor que nuestro 'objetivo' de 10^4."

"Esa es precisamente la idea", dije. "Vamos postulando diversas componentes, las utilizamos para generar otras nuevas y continuamos hasta alcanzar una contradicción; fundamentalmente, que N se ve obligado a ser mayor que el objetivo. Examinadas todas las posibles componentes, hemos terminado."

Parece lógico abreviar los cálculos lo más posible, y de aquí algunas observaciones adicionales que pudieran ayudarnos a ello:

– Al postular una componente podemos suponer que su exponente es una unidad menor que un número primo. Es decir, solamente es preciso considerar los exponentes 1, 2, 4, 6, 10, 12, 16...

– En particular, podemos suponer que el primo especial solamente figura elevado a la primera potencia.

– Si q^{2a} es una componente que aparece elevada a potencia par, esto es, si q no es el primo especial, para un objetivo de 10^4 suponemos que $q^{2a} \leq 10^2$. En general, podemos suponer que q^{2a} es menor o igual que la raíz cuadrada del objetivo.

(*¿Por qué* podemos hacer estas suposiciones? Véanse las soluciones, al final del artículo.)

Consiguientemente, en el caso de un objetivo de 10^4, las únicas componentes q^{2a} que es preciso considerar son 3^2, 3^4, 5^2 y 7^2. (Por ejemplo, $11^2 = 121$ y $5^4 = 625$ son ambos mayores que $10^2 = 100$.) Consideremos por turno las cuatro posibilidades:

3^2: N es divisible por $\sigma(3^2) = 13$. Dado que 13^2 es demasiado grande para ser una componente, el primo especial tiene que ser 13. Entonces $2N$ es divisible por $\sigma(13) = 14 = 2 \times 7$, y, por tanto, 7 es una componente de N. Podemos suponer que su exponente es 2. En tal caso, N es divisible por $\sigma(7^2) = 57 = 3 \times 19$. En particular es divisible por 19. Ahora bien, no es posible que el primo especial sea 19 (hemos visto ya que era 13, pero de todos modos, 19 no es de la forma $4m + 1$, así que *jamás* podrá ser el primo especial) por lo que habrá una componente 19^2, o 19^4... Pero éstas son demasiado grandes y hemos alcanzado una contradicción.

3^4: N es divisible por $\sigma(3^4) = 121 = 11^2$. Hay, pues, una componente 11^2, 11^4,... Pero éstas son demasiado grandes; de ahí la contradicción.

5^2: N es divisible por $\sigma(5^2) = 31$. No es posible que 31 sea el primo especial, pues no es de la forma $4m + 1$, por lo que tiene que haber una componente 31^2, 31^4,... Pero éstas son demasiado grandes. Contradicción.

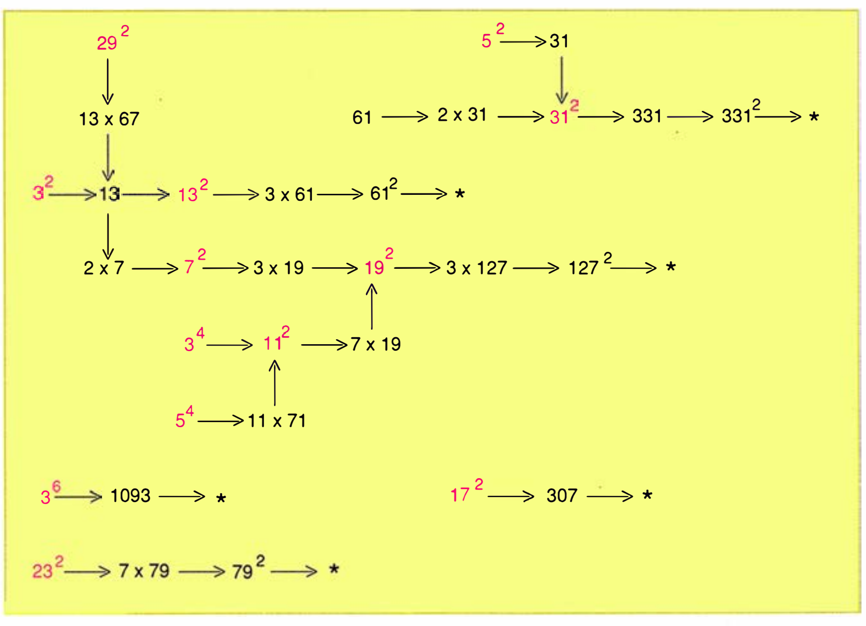
7^2: Hemos eliminado ya la posible componente 7^2 al tratar el caso de 3^2.

Así concluye la demostración.

Fijémonos que, en vez de considerar los 5000 posibles números impares N menores que 10^4, hemos demostrado el resultado examinando ¡tan sólo *cuatro* casos!

El razonamiento entero puede quedar gráficamente resumido como en la figura 2. Los lectores podrían intentar por sí mismos el razonamiento correspondiente al objetivo 10^6: más adelante, en la sección de *Soluciones*, se ofrece una representación gráfica de la mía. He aquí dos indicaciones útiles:

1. La formación de $\sigma(q^{2a})$ correspondiente a una componente q^{2a} suele conducir a *varios* factores primos nuevos; si al menos uno de ellos origina una contradicción, hemos terminado el caso. Por ejemplo, al considerar $\sigma(7^2)$ obtuvimos dos nuevos primos, el 3 y el 19. Del 19 se sigue directamente una contradicción, por lo que no necesitamos examinar también el 3. No obstante, a veces conviene experimentar para hallar el factor más idóneo.



3. Arbol de factores de una demostración de inexistencia de un número perfecto impar menor que 10^6. Los símbolos son los mismos que en la figura 2.

C: Historial de los números perfectos

Fechas	Acontecimientos
c. 300 a. de C.	Euclides demuestra que $2^{p-1}(2^p - 1)$ es perfecto siempre que $M_p = 2^{p-1}$ sea primo.
100 d. de C.	Nicómaco señala que 6, 28, 496, 8128 son los únicos números perfectos comprendidos entre 1 y 10 000 y observa que van alternativamente acabando en 6, 8.
hacia 300	Yámblico afirma erróneamente que la regularidad observada por Nicómaco prosigue indefinidamente.
1202	Leonardo de Pisa afirma la existencia de una infinidad de números perfectos de la forma euclídea.
1456	Un códex anónimo (lat. Monac. 14 908) anuncia el quinto número perfecto, 33 550 336.
1588	Cataldi afirma que M_{17} y M_{19} son primos.
1603	Cataldi afirma que M_p es primo para p = 2, 3, 5, 7, 13, 17, 19, 23, 29, 31, 37.
1638	Descartes cree poder demostrar que todos los números perfectos pares son del tipo euclídeo y que todos los impares son de la forma ps^2, siendo p número primo.
1644	Mersenne enuncia que los once primeros números primos están dados por la fórmula de Euclides, con p = 2, 3, 5, 7, 13, 17, 19, 31, 67, 127, 257.
1722	Euler demuestra que M_{31} es el primero.
1738	Euler da una demostración de que los únicos números primos menores que 100 para los que M_p es primo son p = 2, 3, 5, 7, 13, 17, 19, 31, 41, 47.
1741	En una carta a Kraft, Euler retira de la lista los números $p = 41$ y $p = 47$.
1844	Lebesgue demuestra que todo número perfecto impar tiene al menos cuatro divisores primos distintos.
1849	Publicación de un artículo póstumo de Euler, donde demuestra que todo número perfecto par obedece a la fórmula de Euclides, y que todo número perfecto impar es de la forma $p^{4k+1}s^2$ donde p es primo.
1869	Landry publica una tabla de todos los divisores de $2^p \pm 1$ correspondientes a $p \leqslant 64$, y en particular, demuestra que M_p es compuesto para p = 43, 47, 53, 59.
1876	Lucas afirma disponer de una demostración de que M_p es primo para p = 31 y 127.
1883	Pervushin demuestra que M_{61} es primo.
1887	Sylvester declara haber demostrado que todo número perfecto impar ha de tener al menos 6 divisores primos distintos.
1888	Sylvester demuestra que un número perfecto impar no divisible por 3 posee al menos 8 divisores primos distintos.
1888	Catalan demuestra que un número perfecto impar no divisible por 3, 5 o 7 tiene al menos 26 divisores primos distintos.
1903	Lazzarini trata de demostrar la inexistencia de números perfectos impares, pero su demostración contiene un error.
1903	Cole demuestra que M_{67} es compuesto.
1911	Powers demuestra que M_{89} es compuesto.
1914	Powers demuestra que M_{107} es primo.
1952-90	Se descubren varios otros primos de Mersenne.
1956	Muskat demuestra que no existen números perfectos impares $\leqslant 10^{18}$.
1957	Kanold demuestra que no existen números perfectos impares $\leqslant 10^{20}$.
1973	Tuckerman demuestra que no existen números perfectos impares $\leqslant 10^{36}$.
1973	Hagis demuestra que no existen números perfectos impares $\leqslant 10^{50}$.
1982	Beck y Najor demuestran que no existe un número triperfecto impar $\leqslant 10^{50}$.
1985	Kishore demuestra que todo número triperfecto impar es divisible por 11 números primos distintos.
1989	Brent y Cohen demuestran que no existe un número perfecto impar $\leqslant 10^{160}$.
1989	Brent y Cohen anuncian una demostración de inexistencia de números perfectos impares $\leqslant 10^{300}$.

2. Ciertos casos dan nacimiento a otros con carácter de subcasos. Por ejemplo, el 3^2 anterior condujo a 7^2. Si al analizar los distintos casos no se están haciendo hipótesis particulares referentes a cuál es el primo especial, podemos liquidar simultáneamente el original y el subcaso. De este modo ahorramos trabajo. Dicho de otro modo, las diversas cadenas de divisores que se están examinando pueden hallarse ligadas entre sí en un *árbol de factores*. (O mejor, quizás en un bosque de factores, porque ordinariamente hay varios árboles y cadenas disjuntos.)

"El método utilizado por R. Brent y G. Cohen para la cota 10^{160} consiste, en lo esencial, en esto mismo", dije. "Pero hubieran tenido que examinar un verdadero montón de casos de atender a todas las posibles componentes $q^{2a} \leqslant 10^{80}$. Simplificaron todavía más el razonamiento."

Brent y Cohen demuestran que, si N es un número perfecto impar y ninguno de los primos 127, 19, 7, 11, 13, 31, 3 y 5 son factores de N, entonces N ha de tener al menos 101 factores primos distintos. El producto de los 101 primeros números primos ronda en torno a 10^{473}, mucho mayor que 10^{160} (y, por tanto, lo mismo ocurre para cualquier otro conjunto de 101 primos distintos); he aquí la contradicción. Nos queda ahora ir eliminando por turno los primos anteriores. (El orden elegido resulta especialmente eficiente para evitar repeticiones del trabajo; fue sugerido empíricamente.)

El cómputo de Brent y Cohen se llevó a cabo en un ordenador VAX, pero diversos primos o candidatos a primos fueron registrados en un fichero aparte, para verificar posteriormente su carácter primo. El máximo de los factores hallados fue 5 956 707 000 538 571 084 106 691 363 703, aparecido en $\sigma(61^{42})$. Lo descubrieron utilizando un nuevo método de descomposición en factores primos, el llamado algoritmo de la curva elíptica, debido a Hendrik Lenstra. El cálculo exigió alrededor de 1500 millones de multiplicaciones módulo $\sigma(61^{42})$, efectuadas en varias estaciones de trabajo SUN. He aquí las palabras con las que describen la idea general:

"Proseguimos con las cadenas individuales de factores, utilizando el primo máximo encontrado en la factorización precedente, hasta que obtuvimos una componente que excedía de 10^{80} o hasta que el producto de todos los números primos generados en la cadena, cuyos exponentes estaban frecuentemente ajustados en sentido

ascendente según la forma de Euler, excedían de 10^{160}... Algunas cadenas podrían concluir prematuramente, por aparición de primos previamente eliminados... Aparecieron muchos números compuestos o probables primos para los que no hubo necesidad de descomposición en factores o de aplicar criterios de primalidad. Una vez comprobado que estos números no poseían factores comunes con los que componían la cadena, eran introducidos en el producto con estos números, con exponente 1. La búsqueda de factores comunes mediante el algoritmo de Euclides es más rápida que la factorización."

Euler se levantó y extrajo un documento de su estantería. Me maravilló la facilidad con que encontró el que deseaba, a pesar de su invidencia. Euler poseía una memoria prodigiosa; ciego ya, efectuó una serie de cálculos extremadamente difíciles sobre el movimiento de la Luna, el único problema del que Newton dijo que le había provocado dolor de cabeza. El documento era una copia de una carta de Pierre de Fermat a Marin Mersenne, datada el 10 de agosto de 1638. "En esta carta", dijo Euler, "el gran Fermat declara poseer 'un método analítico merced al cual puede resolver todas las cuestiones atinentes a partes alícuotas, a excepción de la comprobación de primalidad de un número, para lo cual no conocía otro procedimiento que ir tanteando si cada número primo menor que la raíz cuadrada era divisor'. Me ha parecido —continuó Euler— que el cálculo que vos acabáis de exponer se fundaba en métodos más perfeccionados para la factorización de números grandes en producto de primos, y para verificar en números grandes lo atinente a su primalidad. Aparte de eso, en poco mejoran a Fermat."

"Así es", dije yo. "El método es antiguo, pero los cálculos requeridos para utilizarlo se tornan enormes de no encontrar métodos más eficientes para la descomposición en factores. Es un campo en el que actualmente se está avanzando rápidamente. Por ejemplo, Brent y Cohen acaban de anunciar que, debido a perfeccionamientos en las técnicas de factorización, pueden ampliar su cota inferior sobre inexistencia de números perfectos impares hasta 10^{300}. La dificultad principal estriba en la factorización de $\sigma(13^{72})$. En ello les ha ayudado H. J. J. Te Riele, que ha hallado el factor 145 009 586 102 490 829 218 552 548 223 336 637."

Euler se retrepó en su asiento. "¡Espléndido! Viajero del tiempo, en vuestra próxima visita habéis de contarme más acerca de estos métodos de factorización..."

"Bueno, no estoy seguro de saber dar con el camino para volver por segunda vez", mentí, por no imaginar las potenciales paradojas que podrían suscitarse si Euler tomara la cuestión en sus manos. "Lo que me recuerda que he de irme", añadí, no deseando verme en un apuro por sus preguntas. Volví a encaramarme en el torsionador de espacio-tiempo, con la mente firmemente concentrada en Ombilicus y el laboratorio de infinormática del Dr. Zebedee J. J. Bunnydew... La imagen de Euler se fue desvaneciendo, mientras me despedía.

"¡Buen viaje tengáis, viajero del tiempo! Quizás algún día podamos hablar más sobre los números perfectos. Habéis hecho revivir en mí el interés por esta cuestión abierta. De hecho, estoy pensando ya en una versión más general, la de los números que dividen exactamente a la suma de sus partes alícuotas, pero no necesariamente con cociente 2, como los números perfectos!"

"Sí, desde luego, también los números múltiplemente perfectos son fascinantes", respondí. "Pero os advierto que plantean cuestiones igualmente difíciles. En 1982, W. Beck y R. Najar demostraron que no existe ningún número triperfecto impar [un número N tal que $\sigma(N) = 3N$] que sea menor que 10^{50} (*El lector podría encontrar de su gusto una adaptación del método de la cadena para tratar números triperfectos impares. Piense en una forma estándar para éstos.*)

Soluciones

1. Sea

$$N = p^a q^b \ldots r^c.$$

Entonces,

$$\sigma(N) = \sigma(p^a)\sigma(q^b)\ldots\sigma(r^c).$$

Pero

$$\sigma(p^a) = 1 + p + p^2 + \ldots + p^a = (p^{a+1} - 1)/(p - 1)$$

Por consiguiente

$$\sigma(N) = (p^{a+1} - 1)\,(q^{b+1} - 1)\ldots(r^{c+1} - 1)/(p - 1)\,(q - 1)\ldots(r - 1).$$

2. Dado un primo cualquiera p, la fórmula anterior para $\sigma(p^a)$ demuestra que, si $a + 1$ divide a $b + 1$, entonces $\sigma(p^a)$ divide a $\sigma(p^b)$, por lo que al formar las cadenas de divisores podemos limitarnos a considerar los casos en que $a + 1$ es primo.

Si q^{2a} es una componente entonces $N \geq q^{2a}\,\sigma(q^{2a}) > q^{4a}$, así que $q^{2a} < \sqrt{N}$.

3. Para una demostración de que ningún número perfecto impar puede ser $> 10^6$, véase la figura 3.

Libros

Sobre progreso y evolución, superconductividad, educación española, mitosis y clásicos de hoy

Antonio Prevosti, Juan Bartolomé, Antonio Domínguez,
Mercedes Durfort y Luis Alonso

EVOLUTIONARY PROGRESS. Dirigido por M. H. Nitecki. The University of Chicago Press; Chicago, 1988.

¿Cómo encaja la pequeña parcela de conocimientos de nuestra especialidad científica en las concepciones generales del mundo en que vivimos? Esta pregunta a veces despierta inquietud en los científicos, como se refleja en el contenido del libro que comentamos a continuación, que trata de las relaciones de la teoría de la evolución biológica con la idea de progreso en la naturaleza. Así lo plantea Matthew H. Nitecki en el prefacio: "En nuestro mundo prevalece el pesimismo y toda la estructura de la sociedad es un desafío a la esperanza. Para Jacques Monod el azar explica el propósito de la vida, Paul Kennedy profetiza que el desarrollo de las grandes potencias conduce inevitablemente a su decadencia, mientras que Gunther Stent concluye que Dios está muerto, el arte está muerto, la música ha llegado a su cénit y la ciencia ha alcanzado sus límites. La desesperación parece ser la moda actual. Se presume que el crecimiento de la población es una bomba mortífera, la escasez de alimento producirá hambruna general, el SIDA diezmará nuestra población, mañana imperará la pobreza, alzamientos religiosos y fundamentalistas destruirán la libertad, el sistema monetario actual se colapsará, se destruirá el orden social, el mercantilismo es omnipresente, las ideas están muertas, la polución nos asfixiará, la falta de combustibles nos congelará y los tártaros encenderán las antorchas de la guerra con megatones.

No obstante, tan sólo para sobrevivir en nuestra vida cotidiana debemos confiar en mañana... Por tanto, la idea de progreso no puede estar muerta, sólo debe de haber cambiado la interpretación de su significado."

Este planteamiento de Nitecki se sale claramente del campo de la evolución biológica. No obstante, la evolución biológica y la de algunos aspectos de la sociedad humana presentan muchos rasgos comunes. Así lo expresa, por ejemplo, Ruse al final de su artículo en este libro diciendo: "..aquí hago mi tercera y última sugerencia, simplemente que la ciencia en general –de la cual la teoría de la evolución es un caso particular– es el área de la actividad humana donde nos parece que conseguimos más inequívocamente progreso... Lo que funciona en el mundo del pensamiento, seguramente también debe de funcionar en el mundo de los organismos."

Evolutionary Progress reúne las contribuciones de un grupo de destacados especialistas en biología evolutiva e historia y filosofía de la ciencia. En su primera parte se analiza el concepto, o mejor dicho los conceptos de progreso, desde el punto de vista filosófico. Según Hull se ha pasado de la idea clásica y antropocéntrica de progreso al concepto de direccionalidad en la evolución biológica y, además, rechaza que ésta tenga validez universal. Provine, partiendo de la posición favorable a la existencia de progreso de Julian Huxley, se muestra contundente contra la aceptación general de esta idea por el neodarwinismo. Ayala hace una disección minuciosa y matizada del concepto de progreso en la evolución biológica. Ruse analiza la idea de progreso a lo largo del pensamiento evolutivo.

En la segunda parte del libro se pasa revista a ciertos aspectos históricos, junto con puntos de vista sobre la idea de progreso situados fuera de la corriente principal de la biología evolutiva. Son cuatro los artículos que integran esta parte. Robert J. Richards examina históricamente los fundamentos morales de la idea en el desarrollo del darwinismo. Robert C. Richardson y Thomas C. Kane la analizan, principalmente, en relación con la ortogénesis. Robert C. Dunnell la estudia en la evolución cultural. Finalmente, Adam Urbanek presenta las ideas desarrolladas en la URRS, principalmente por el especialista en anatomía comparada y biólogo evolutivo A. N. Severtsov, sobre la aplicación de la idea de progreso a la evolución morfofisiológica.

Los autores de la última parte consideran, desde distintos puntos de vista, las posibilidades de llegar empíricamente a un concepto de progreso con validez científica. Maynard Smith analiza las razones para desconfiar de la idea de progreso; pero, al tratar de los distintos niveles en los que actúa la selección, moléculas, células, organismos, grupos y especies, opina que el paso de un nivel a otro de orden superior es una revolución que conduce a la individualidad en el nivel superior, con supresión de los efectos disruptivos de la selección en el inferior, lo cual es, en cierto sentido, progreso. Wimsatt y Shank examinan varios mecanismos microevolutivos, es decir, de genética de poblaciones, que a largo plazo pueden determinar el aumento del número de genes mantenidos por la selección, lo que hace posible, en su opinión, explicar cambios macroevolutivos. Opinan que, si su interpretación es correcta, estos mecanismos pueden constituir la base para reintroducir la noción de progreso en el pensamiento evolutivo.

Según Wiley la tendencia al incremento de organización observada en el curso de la evolución está ligada al aumento de entropía presente en todos los sistemas termodinámicos, tanto físicos como biológicos, y concluye que el aumento de entropía es antitético al concepto de progreso del siglo XIX. Raup se plantea si los datos de la paleontología permiten deducir que hay progreso. Opina que pocos observadores, al comparar los biota modernos con la biología del Precámbrico, negarían que ha ocurrido algún tipo de progreso, pero cree que la consideración de lapsos de tiempo más cortos hace difícil confirmar esta conclusión. Partiendo, principalmente, de los datos sobre el aumento de supervivencia de los géneros a lo largo del tiempo, llega a la conclusión de que es necesario disponer de una escala más precisa del tiempo geológico y de medios más eficaces para substraerse a la atracción de lo reciente, para llegar a una conclusión más segura. Fundado también en el estudio de los fósiles, Gould rechaza rotundamente la idea de progreso, pero considera legítimo desde el punto de vista científico admitir direccionalidad en la evolución.

La conclusión más general a la que se llega al leer el libro es que hay muy

poco acuerdo sobre la existencia de progreso en la evolución y sobre el contenido de la idea de progreso. Estos dos aspectos están en parte relacionados. Muchos científicos y filósofos de la ciencia rechazan la idea de progreso por el origen histórico que tiene en la teoría de la evolución. Estaba relacionada con la concepción antropocéntrica del universo, vigente cuando se desarrolló esta teoría. El hombre era la finalidad del universo y al substituirse la concepción estática del mundo por la dinámica, basada en la evolución, fueron muchos los que pensaron que ésta tenía una meta, la aparición del hombre. Además, de acuerdo con las ideas de la época, el progreso evolutivo no era una característica intrínseca del universo, sino que estaba determinado desde fuera por el Creador. Esta idea de progreso no es científica y, por ser este su sentido primitivo, algunos autores como Hull, Provine y, en parte Gould, la toman como la única válida, lo que les lleva a rechazar que la evolución sea progresiva. En relación con esta cuestión es muy aclaratorio el artículo de Ayala, que analiza el concepto de progreso y lo define en términos generales, sin ligarlo al antropocentrismo histórico.

Otros autores rechazan, con mayor o menor decisión, que se manifieste progreso general en la evolución biológica, pero al mismo tiempo admiten, también en distintos grados, que pueda aparecer en aspectos parciales. Así, su actitud resulta dubitativa. Esto se refleja en el artículo de Maynard-Smith. Ayala se expresa más claramente, afirmando que no existe un criterio único que permita afirmar que en la evolución biológica hay progreso, pero admite que con criterios distintos cabe aceptarlo en diferentes líneas filogenéticas. No obstante, estima que la idea de progreso no es aceptable como concepto científico. En primer lugar, y fundamentalmente, porque la noción de progreso es axiológica, implica un juicio de valor, es subjetiva y, por tanto, no científica. En segundo lugar, porque considera fallidos los intentos de definir el progreso como un concepto biológico. Las conclusiones de Ayala no son dubitativas, pero reflejan la complejidad con que ve la situación, cuando después de manifestar que es dudoso que se gane nada hablando de progreso en la evolución, dice que le gustaría que el término progreso se eliminara totalmente del discurso científico. No obstante, luego añade que le parece improbable que esto ocurra, porque es una noción que parece estar irremediablemente arraigada entre las categorías del pensamiento del hombre moderno y es probable que se continúe usando en relación con la evolución.

Muy significativo respecto al vínculo de la noción de progreso con el pensamiento evolutivo es el artículo de Ruse. Para este autor, desde Darwin la idea de evolución va asociada a la de progreso, en el pensamiento de casi todos los biólogos evolucionistas. Que Darwin era 'progresionista' lo afirma rotundamente diciendo: "los que niegan que Darwin era 'progresionista' –y yo era uno de ellos– están completamente equivocados". Después indica que la idea de progreso está clara en el pensamiento de los principales artífices del neodarwinismo, Julian Huxley, Dobzhansky, Mayr, Simpson, Stebbins, etc. Por último, analiza la obra de destacados evolucionistas modernos detractores de esta idea, como E. O. Wilson, Dawkins y Gould, y llega a la conclusión de que también está impregnada de la idea de progreso.

En realidad, si el concepto de evolución se diferencia en algo del de simple cambio es, seguramente, en implicar la idea de progreso en alguna forma u otra. No obstante, hay que precisar que, en relación con la posición del hombre, esto no significa que no se reconozcan sus imperfecciones. De hecho ya se hacía así en las ideas preevolutivas. Como especifica Maynard-Smith en su contribución a este libro, en la antigua escala de los seres se situaban, por encima del hombre, las distintas categorías de ángeles. (A. P.)

HIGH TEMPERATURE SUPERCONDUCTIVITY. Dirigido por Jeffrey. W. Lynn. Springer Verlag; Berlín, 1990.

La superconductividad de alta temperatura crítica (SATC) es un fenómeno físico cuya posibilidad había sido "anatemizada" durante décadas. Su descubrimiento en 1986 condujo a una proliferación de artículos técnicos y de divulgación solamente comparable a la teoría de relatividad, salvando las distancias. Las palabras de K.A. Muller, premio Nobel por el descubrimiento del fenómeno, "estoy deseando que se encuentre otro fenómeno pronto para que se empiece a estudiar la superconductividad de alta temperatura crítica en serio", nos indican hasta qué punto se ha producido un desbordamiento de producción "rápida" pero superficial.

La superconducción había dejado de ser un tema de interés científico por los escasos avances conseguidos en los años setenta; prácticamente se había abandonado. Al suscitarse el interés universal por la SATC hubo que revisar los conocimientos existentes, teóricos y experimentales, sobre la superconducción en materiales convencionales. El libro que nos concierne consiste en un conjunto de trabajos que fueron presentados en un curso de doctorado de la Universidad de Maryland en respuesta a dicha necesidad, que se celebró en 1987.

Como toda obra de este tipo, en que se yuxtaponen trabajos de distintos autores, son frecuentes las reiteraciones y, en algunos casos, cierta confusión en la terminología.

El capítulo 1 tiene un carácter básico. Describe el fenómeno de la superconductividad e introduce la terminología de los modelos de London, Ginzburg y Landau, Bardeen, Cooper y Schrieffer, para definir los parámetros de dimensión relevantes en la clasificación de los superconductores como tipos I y II. Por necesidad deja muchos puntos sin demostrar, pero asienta razonablemente los puntos nucleares de dichas teorías. También es útil como vademécum de las expresiones y valores de constantes frecuentemente utilizadas en un primer momento en la caracterización fenomenológica de los nuevos SATC. Teniendo en cuenta el objetivo didáctico declarado por el editor, que también es editor de este capítulo, hemos de consignar que adolece de algún error en las figuras y, sobre todo, de una falta completa de otras que hubieran sido esenciales para facilitar la comprensión del texto, en particular al pasar revista a los conceptos de longitud de penetración, correlación y estado mezcla.

Le sigue un capítulo fundamentalmente teórico, donde se presentan las teorías de superconductores convencionales aceptadas en la actualidad, como la justificación de la presencia de vórtices en los superconductores tipo II y su estructura ("red de vórtices de Shubnikov"). Realmente interesante es el análisis crítico del efecto de los mecanismos de interacción frente al desorden (estructural, causado por defectos u otros). Ofrece algunas predicciones todavía sin verificación experimental. Inevitablemente, su autor profundiza en aquellos temas en que ha trabajado de forma directa, pero ello no le resta calidad.

Al exponer el efecto Johnson en sus diversos casos, se pone a punto la herramienta teórica para tratar posteriormente el problema de las interacciones entre los granos de los SATC cerámicos, así como en el caso de las láminas delgadas granulares. Se detiene en el ruido de origen cuántico

debido a fluctuaciones térmicas, efecto pocas veces considerado.

El capítulo cuarto trata de las estructuras cristalográficas de las familias de SATC más importantes encontradas hasta mediados del 88 con un esquema lógico basado en la estructura parental de perovskita cúbica. Es una buena referencia de partida, ya que se ha depurado bastante la información estructural desde esa fecha. Difícil de entender es el dominio de las propiedades electrónicas de los SATC. Se da aquí una revisión profunda de las predicciones de teoría de bandas y su comparación con resultados experimentales. Llama la atención la sugerencia de que las cadenas CuO en $YBa_2Cu_3O_{7-\delta}$, o las capas BiO o TlO en los compuestos basados en el Bi y en el Tl, tienen carácter metálico e influyen en las propiedades superconductoras de los materiales.

Cuando se describen los métodos de síntesis de preparación de los SATC más sencillos y los métodos de primera caracterización, se advierte el paso del tiempo desde que fue escrita la obra. No se citan algunas de las vías de preparación mediante precursores organometálicos, ya conocidas entonces. Tampoco se explica apenas la preparación de capas finas, que constituye uno de los campos donde se puede prever más desarrollo en aplicaciones tecnológicas de estos materiales.

El estudio de las propiedades macroscópicas de los SATC parte del comportamiento correspondiente en los superconductores convencionales, fermiones pesados o de ondas de densidad de carga o espín. Es un capítulo excelente, inteligible para un recién licenciado con conocimientos de estado sólido, a diferencia de algunos de los otros capítulos. Por otra parte, la mayoría de las cuestiones dejadas abiertas en el texto no han recibido respuesta definitiva todavía (modelo de acoplamiento fuerte BCS frente al de enlace de valencia resonante, modelo bipolarónico o partículas de fase fraccionaria).

Al revisar el comportamiento magnético microscópico de las familias de compuestos asociados a los SATC, se demuestra el carácter bidimensional magnético de las fluctuaciones de los espines asociados a las posiciones de metal de transición (Cu, Fe). Y queda probado el aislamiento electrónico de los metales sustituyentes en posiciones distintas de los átomos de Cu. Se introduce luego la teoría BCS (Bardeen, Cooper y Schrieffer) en formulación de segunda cuantificación y se discute en profundidad los límites inherentes a la hipótesis de acoplamiento débil.

En el último capítulo se describen algunas aplicaciones en dispositivos de materiales superconductores y un análisis comparativo de las figuras de mérito que los nuevos SATC imponen frente a los superconductores convencionales. Cuatro años después del descubrimiento de los SATC se cree que en el campo de las aplicaciones en sensores y dispositivos electrónicos pueden encontrarse sus aplicaciones más inmediatas, sobre todo una vez que se empieza a dominar la deposición de capa delgada. Por ello este capítulo es francamente ilustrativo del potencial y limitaciones de esta tecnología. Declara el editor en la introducción que comenzó organizando un curso de iniciación a los SATC, de orientación didáctica. Sin embargo, el resultado es un libro de consulta valioso, pero de lectura sólo apta para el investigador especializado. (J. B.)

LA ESCUELA EN EL CATASTRO DE ENSENADA, por Carmen Labrador Herraiz. **EDUCACION Y ECONOMIA EN LA ESPAÑA ILUSTRADA**, por Agustín Escolano Benito. **PABLO MONTESINO. CURSO DE EDUCACION**, por Anastasio Martínez Navarro. **PASCUAL VALLEJO, ILUSTRADO Y REFORMADOR DE LOS ESTUDIOS**, por Carmen Labrador Herraiz. Ministerio de Educación; Madrid, 1988.

Estos cuatro volúmenes, editados en 1988, forman parte de la contribución aportada por el Ministerio de Educación y Ciencia a los actos organizados con motivo del II centenario de la muerte de Carlos III, ocasión aprovechada por numerosos organismos oficiales y particulares para fomentar el conocimiento de la Ilustración española y del Soberano que, juntamente con sus ministros, propiciaron un decisivo avance de la vida española en todos los órdenes. Los cuatro forman parte de una serie sobre la Educación y la Ilustración, que comprende estudios originales y reproducción de antiguos textos valiosos cuya consulta resultaba muy difícil por la rareza de los ejemplares existentes.

El volumen titulado *Educación y Economía en la España Ilustrada* detalla la creación y funcionamiento de una serie de centros educativos que tendían a remediar las deficiencias de que adolecían las enseñanzas técnicas. Las enseñanzas tradicionales que se impartían en las universidades formaban teólogos y juristas, pero no economistas, ingenieros, marinos y otros profesionales de los que España tenía gran necesidad. Campomanes, en sus *Discursos sobre la Industria Popular* y otros escritos, fue el abanderado de esta tarea en pro de una enseñanza que se apartara de los moldes teóricos y rutinarios en que estaba inmersa. Especial interés tenían las enseñanzas náuticas para la nación que poseía el más vasto imperio colonial conocido hasta entonces. Para responder a esta necesidad surgieron el Colegio de San Telmo de Sevilla, la Escuela de Guardias Marinas de Cádiz, el Observatorio Astronómico de San Fernando, las escuelas de Náutica de Bilbao y Barcelona y otros centros en los que se fomentó el estudio de las matemáticas modernas y de la física de Newton, tratando de recuperar el notable retraso que en este punto padecía España.

La monografía sobre Pascual Vallejo elaborada por Carmen Labrador nos restituye la imagen de un olvidado pedagogo que en 1791 escribió un *Discurso sobre la necesidad de una reforma general de los métodos de educación.* El Discurso, que ahora se imprime por primera vez, critica los vigentes métodos de enseñanza universitaria y sugiere las oportunas reformas de acuerdo con el ideario de la Ilustración.

Otro texto recuperado es el *Curso de Educación* de Pablo Montesino, figura que se escapa del marco de la Ilustración, pues nació en 1781 y desarrolló su actividad en la primera mitad del siglo XIX. En el texto que ahora se edita Montesino resumió las lecciones impartidas en el Seminario Central de Maestros, germen de las futuras Escuelas Normales, y contienen lo que él consideraba "el bagage pedagógico mínimo e imprescindible para los nuevos maestros que habían de acometer la reforma educativa".

Nos detendremos algo más en el examen de *La Escuela en el Catastro de Ensenada* porque es una aportación muy estimable al conocimiento de la realidad escolar en el siglo XVIII y ha exigido de su autora laboriosas investigaciones. Como es sabido, el Catastro elaborado por orden del marqués de la Ensenada entre 1748 y 1752 fue un esfuerzo gigantesco, la más grande y minuciosa encuesta realizada en los reinos de Castilla, base inagotable para multitud de estudios y monografías.

El volumen de referencia contiene ante todo una idea general de dicho Catastro; después, un esquema del panorama social de la provincia de Guadalajara en la época en que fue confeccionado; un estudio detallado de la situación de la enseñanza y los enseñantes tal como se desprende del mismo; una comparación con los datos que suministra el Catastro acerca del número y retribución de los maestros y preceptores en los demás reinos y provincias de España y, finalmente,

unas copiosas tablas estadísticas que resumen todos los datos demográficos y socioeconómicos que de dicha fuente pueden extraerse acerca de cada uno de los municipios de Guadalajara: número de vecinos, distribución por edades y sexos, profesionales, etc. Se trata de una investigación laboriosa que, sin duda, habrá costado no pocos años y esfuerzos a la autora.

En lo relativo a la enseñanza primaria los datos que suministra confirman y corroboran con numerosos ejemplos lo que por otras fuentes ya sabíamos: la escasez de escuelas y maestros, su mala calidad, su miserable salario, el escaso interés de las autoridades y un muy largo etcétera. Los *ilustrados* españoles dieron pruebas de gran preocupación por los problemas de la enseñanza, pero se fijaron, sobre todo, en la cúspide, en las universidades; no se dieron cuenta, hasta muy tarde, de que mientras fallara la base no podría construirse un edificio consistente; hay que esperar a los finales del reinado de Carlos III para hallar unas disposiciones legislativas de tipo general que, como tantas otras reformas, tuvieron poca efectividad por falta de medios económicos; el Estado no consignaba en sus presupuestos ninguna cantidad a estas atenciones; desviaba la carga hacia los municipios, y éstos, salvo excepciones, no estaban dispuestos a hacer sacrificios por la enseñanza primaria.

De aquí el panorama lamentable que se dibuja en estas páginas. No hay en el Catastro una estadística de escuelas sino de maestros; más indirectamente se deduce que en los pueblos en que no había maestros no podía haber escuelas; éste era el caso de muchas pequeñas entidades de población. A los maestros no se les exigía todavía ningún título, ninguna garantía de solvencia profesional. La enseñanza que dispensaban se limitaba a leer, escribir, conocer las cuatro reglas de aritmética y unos rudimentos de Doctrina Cristiana. Los emolumentos que percibían unas veces iban a cargo del municipio, otras de los padres de los alumnos o bien de ambos conceptos, y siempre se mantenían en unos límites modestísimos. Como la finalidad del Catastro era averiguar los ingresos de cada habitante para instaurar un nuevo sistema impositivo, conocemos estas remuneraciones con todo detalle; predominaba el pago en especie, sobre todo en trigo: ocho fanegas, diez fanegas, más algunos reales de vellón y tal vez alguna ventaja adicional, como el pago de la casa.

Los resúmenes catastrales de la provincia de Guadalajara ponen de relieve esta insuficiencia: mientras los 40 médicos existentes en la provincia ganaban en total 192.559 reales, o sea, 4.810 reales anuales cada uno, y los 116 sacristanes 111.000, casi a mil reales individuales, los 116 maestros de primeras letras ingresaban sólo 47.512, o sea, cuatrocientos reales. ¡Poco más de un real diario! Como era imposible subsistir con semejante paga, era frecuente la doble o triple profesionalidad; el maestro solía ser a la vez el sacristán de la parroquia, y con frecuencia también el escribano o el alguacil. Podría objetarse que dada la exigüidad demográfica y económica de la mayoría de los municipios guadalajareños no se trata de un caso representativo, pero la amplia muestra que ofrece la autora sobre otros municipios españoles confirma que en otras regiones más favorecidas la situación no era mucho mejor. A falta de datos sobre Madrid, la ciudad donde el estamento docente acumulaba mayores ingresos era Cádiz: 201.718 reales; pero Sevilla sólo figura con 110.575, Jerez con 17.000, Málaga 15.400, Jaén 5882...

Evidentemente, hay que contar con el suplemento representado por establecimientos religiosos, no reseñados en el Catastro, con la enseñanza recibida en el seno de la familia y otras fuentes de alfabetización, pero aún así el resultado final se reflejaba en tasas de analfabetismo enormemente elevadas; más, por supuesto, en el campo que en las ciudades, y mucho más en el sexo femenino que en el masculino. Y lo más lamentable es que estas deficiencias sólo disminuyeron en corta medida y muy lentamente en el siglo XIX. (A. D.)

MITOSIS, dirigido por J. S. Hyams y B. R. Brinkley. Academic Press; Londres, 1990.

"El mayor avance que puede concebirse en biología sería el descubrimiento de la inestabilidad que conduce a la división continua de la célula. Cuando veo cómo se divide una célula, siento lo que sentiría un astrónomo si pudiera ser espectador de la formación de una estrella doble: que ante mis ojos se está llevando a cabo un acto original de creación", así escribía en 1906 Bateson.

Desde los trabajos de Flemming sobre el núcleo de células animales y los de Strasburger en células vegetales, los denominados cromosomas por Waldeyer han sido motivo de continuados estudios y se han revelado como estructuras particularmente difíciles de interpretar al observarlas en imágenes ultraestructurales con las técnicas habituales de la microscopía electrónica de transmisión. La división celular observada con un microscopio de contraste de fases ha sido, es y será uno de los procesos más espectaculares y fascinantes de la citología y de la biología en general.

El tratamiento de los tejidos con el alcaloide extraído del *Colchicum,* la colchicina, facilitó el reconocimiento y recuento de los cromosomas al quedar la mitosis bloqueada en metafase, etapa en que las cromátidas tienen mayor volumen y, por consiguiente, son más fáciles de visualizar tras la aplicación de colorantes como la orceína o la hematoxilina férrica, o mediante impregnación argéntica. Hasta el estudio ultraestructural de la mitosis no se llegó a comprender el motivo del aparente aumento de su número en los tejidos sometidos a la acción de la colchicina; finalmente se pudo comprobar que ésta inhibía la formación de microtúbulos fusoriales impidiendo el ensamblaje de las tubulinas y bloqueando así la dinámica cromosómica.

Posteriormente se han hallado otros alcaloides, como el taxol, extraído del tejo, que estabiliza los microtúbulos durante el procesado de la fijación química de las muestras biológicas, impidiendo que los microtúbulos se desvanezcan y posibilitando el estudio utraestructural de los implicados en la mitosis. Dado el importante papel que desempeñan los microtúbulos en dicho proceso de división de las células animales y vegetales, cualquier avance en el conocimiento de los microtúbulos y del citoesqueleto representa disponer de más datos para comprender mejor la dinámica de los cromosomas en el proceso mitótico.

Coincidiendo con el centenario del hallazgo de la acción mitoclástica o inhibidora de la colchicina por Pernice aparece esta monografía sobre la mitosis, justo en el año en que la prestigiosa revista *Cell* publica un hallazgo realmente importante y que aunque era sospechado desde hacía décadas no se había logrado evidenciar: Hall y otros, aplicando técnicas de hibridación in situ, ponen de manifiesto la presencia de ADN en los corpúsculos basales de los flagelos de *Chlamydomonas.* Recordemos que los corpúsculos basales se originan durante la ciliogénesis a partir de los centríolos, por lo que la dinámica y mecanismo molecular de la génesis y funcionamiento de estos orgánulos microtubulares empieza a vislumbrarse, saliendo del campo de las hipótesis para pasar al de la confirmación de las hipótesis.

Mitosis reúne diez trabajos sobre las estructuras implicadas en el proceso de la división celular, así como

sobre los factores bioquímicos que lo catalizan. En cada capítulo, de autoría distinta, hay una introducción histórica al tema, cuya fecha de inicio, variable, depende del tema. Sigue una descripción metodológica bastante pormenorizada, así como el tratamiento estructural, ultraestructural y bioquímico, acompañado siempre de una iconografía escogida que, a pesar de la calidad mediocre del papel, permite apreciar los detalles señalados. Esquemas tridimensionales ponen de relieve las hipótesis interpretativas de los componentes del aparato mitótico. Cada capítulo culmina con una nutrida relación bibliográfica. Encontramos a faltar algún capítulo dedicado a las mitosis no convencionales, es decir, a las mitosis acéntricas de los protozoos; aunque hay alguna referencia a la de algas unicelulares del grupo de las diatomeas, creemos que es insuficiente.

La obra se inicia con un trabajo de Kent McDonald, de la Universidad de Colorado, sobre el huso acromático. Altamente didáctica, constituye un capítulo básico para conseguir el desarrollo de los sucesivos. Vandre y Borisy, de la Universidad de Wisconsin, abordan el ciclo centrosómico en células animales. Brinkley y colaboradores, de la Universidad de Alabama, se ocupan del cinetocoro. ¿Es el centrómero la misma estructura que el cinetocoro? Su estudio está ilustrado por numerosos esquemas tridimensionales, inéditos e informativos. A pesar de los años transcurridos desde que Schrader definiese en 1953 el cinetocoro como un elemento esencial en la motilidad de los cromosomas, esta región, con propiedades tintoriales diferenciales del resto de la estructura y con una disposición ultraestructural en placa trilaminar específica, continúa siendo un enigma.

"La dinámica de los microtúbulos y el movimiento de los cromosomas" es el capítulo que nos introduce en los aspectos más sustantivos de la mitosis. Mediante la utilización de diversas drogas, aplicadas a distintas concentraciones, se llega a la formulación de la atrayente hipótesis de trabajo del "pac-man, polar ejection" para explicar el movimiento de los cromosomas, teoría actualmente admitida por la mayoría de autores. Salmon, de la Universidad de Carolina del Norte, aporta los datos moleculares y bibliográficos de la tubulina en la formación de los microtúbulos y la relación de éstos con el movimiento de los cromosomas.

Bloom y Vallee, de la Universidad de Texas, estudian las proteínas asociadas a microtúbulos en el caso del huso acromático de los huevos del erizo de mar. Se explayan en el papel estabilizador del taxol en el ensamblaje de los microtúbulos, para lograr la identificación ultraestructural y bioquímica de las proteínas asociadas a los mismos. Presentan bandas electroforéticas de alta resolución, en presencia y ausencia de taxol; se puede hacer un estudio comparativo de dichas proteínas asociadas. Junto con los inmunoblots hay interesantes imágenes obtenidas por immunofluorescencia de huevos tratados con anticuerpos monoclonales.

No podía faltar un capítulo dedicado a la calmodulina en la funcionalidad del huso acromático. Wesch y Sheet, de la Universidad de Michigan, son los autores del capítulo a ello dedicado, con espléndidas imágenes de la localización por inmunofluorescencia de la calmodulina. Hay también otras ultraestructurales del inmunomarcaje de la misma, mereciendo destacarse su ubicación en los puentes intercelulares que se forman en la telofase y que pueden matenerse para crear cistes o poliplastos; estas formaciones abundan en los primeros estadios de la gametogénesis y durante las primeras etapas de la eritropoyesis.

Hepler nos introduce luego en las membranas del aparato mitótico, aunque se muestra algo superficial a propósito de la dinámica del retículo endoplasmático, del complejo de Golgi y también del plasmalema.

De la estructura y función del huso nos habla Cabral, de la Universidad de Texas, quien se centra en los aspectos morofológicos y funcionales del huso acromático que se obtiene al trabajar con diversos mutantes de levaduras.

Tras el trabajo sobre mitosis in vitro, de Zacheus Cande, de la Universidad de California, McIntosch cierra la monografía con una mirada hacia el futuro: "Mitosis: perspectivas". Recopilando lo expuesto por quienes le preceden, destaca el interés de la mitosis para acometer modelos celulares "sencillos", es decir, levaduras y hongos. La obra, en su conjunto, es una puesta al día, en sus distintos planos (molecular, estructural y ultraestructural), del proceso mitótico. Si bien esta información es cuantiosa, queda aún por hallar la respuesta a la inevitable pregunta formulada desde principios de siglo: ¿cuáles son los factores que determinan que una célula inicie y culmine su proceso de división? (M. D.)

MOLECULAR BIOLOGY. A SELECTION OF PAPERS. Compilación de S. Brenner. Academic Press; Londres, 1989. **IMMUNOLOGY 1930-1980.** Dirigido por Pauline M. H. Mazumidar. Wall and Thompson; Toronto, 1989.

No es fácil historiar el desarrollo reciente de la ciencia. Exige, en quien lo acomete, poseer un notable grado de profundización en la misma, si quiere evitar la fantasía, la falacia por consiguiente, de reducirla a las grandes fechas o a los grandes hitos que, como novedad, vemos hoy incluidos en las introducciones a muchos manuales universitarios. Además, habida cuenta del progreso de las distintas disciplinas, resulta punto menos que imposible no desfigurar el marco general cuando se contempla desde la particular atalaya donde uno se encuentre mejor instalado.

La dificultad no exime, empero, de su necesidad. Para el propio avance de la ciencia precisamos conocer su historia, quizá no estrictamente la tejida de fechas, personas y lugares, pero sí la trenzada de ideas que explican por qué estamos abordando ahora determinado problema y lo hacemos de esta manera con tales medios. Esta aproximación puede hacerse desde múltiples ópticas. Dos no son las últimas en importancia: conocer el trabajo de quienes nos precedieron, con sus propias palabras y titubeos, y oír la reflexión de quienes hicieron el presente de la ciencia con medios del pasado. Sydney Brenner nos facilita la primera tarea en *Molecular Biology. A Selection of Papers*. Los ensayos-recuerdo de la historia de la inmunología recogidos en *Immunology 1930-1980* pertenecen al segundo grupo.

Entre los placeres que yo recomendaría a los lectores de la revista, sean o no biólogos, tiene la compilación de Brenner un altar. La cosa es muy sencilla: en abril de 1959 aparece la revista *Journal of Molecular Biology*, con el propósito de dar cuerpo doctrinal a una disciplina emergente, cuyos seguidores podían contarse con los dedos de la mano. Tras la dirección de John Kendrew, se hizo cargo de ella Brenner, quien al cumplirse los treinta años de su aparición vuelve a publicar aquí lo mejor de los cien primeros números (o sea, hasta 1975).

Podríamos ponderarle al lector que lo que aquí se resalta constituye hoy la sucesión de capítulos de cualquier texto de biología molecular; es decir, se ha convertido en doctrina recibida o enseñanza oficial, lo mismo en genética que en bioquímica. Y así, a modo de ejemplo, lo que los iniciados abrevian coloquialmente como "pajamo" apareció en el primer número, cuyo contenido incoa la introducción a la biología molecular estándar: la

importancia de la represión y la síntesis de un factor citoplasmático. Se trata de un artículo archifamoso e inaccesible para la mayoría de los alumnos y profesores: "The Genetic Control and Cytoplasmatic Expression of 'Inducibility' in the Synthesis of beta-Galactasidase by *E. coli*", firmado por Arthur B. Pardee, François Jacob y Jacques Monod.

Hay artículos doctrinales (interés de los profagos, mecanismos reguladores, sistemas de control, proteínas alostéricas, función del ARN de transferencia –en un comienzo llamado ARNs–, reconocimiento codón-anticodón, etcétera). Los hay de innovaciones técnicas, que aceleran el desarrollo al facilitar el ensayo en determinado frente atascado; nos referimos a la clonación bacteriana, enzimas de restricción o la secuenciación nucleotídica por hibridación de Southern.

Podríamos entreternos en los artículos, uno por uno. Merece la pena. Pero lo que me interesa destacar aquí no es tanto el valor intrínseco de lo aportado, cuanto el método. Quizás en él esté la razón del éxito de esos logros que hoy consideramos históricos. Por actuar con método entiendo, entienden los firmantes, comenzar por exponer con nitidez la situación y el problema (relatar la coherencia doctrinal interna de lo sabido y definir dónde está la aporía a deshacer); en la mostración de protocolos, puntillismo y sobriedad; en los experimentos, recuento de los pasos seguidos con todas las ayudas de microscopía o interpretación de cromatogramas demandados; en el análisis ("discussion"), los fallos de los ensayos realizados por otros, los caminos que en sus tanteos abrieron, las aclaraciones que nuestro trabajo aporta con sus limitaciones.

Lo que es texto objetivo se convierte en recuerdo personal a propósito del desarrollo de la inmunología en la reunión mantenida en Toronto en julio de 1986, cuyas conferencias constituyen el segundo libro de reseña. También en este volumen hay un exudado que me parece digno de atención: la importancia de los "factores secundarios" en el desarrollo de la ciencia. Son éstos, entre otros, el doctorando que acomete la experimentación de la idea del director, la charla de sobremasa, la cesión de cepas o de muestras de dudosa por ignorada naturaleza, etcétera, que no suelen recogerse en la estampación del artículo. El doctorando puede firmar el trabajo, pero no se dice si confirma la idea de partida o la remodela; la charla de café puede aparecer perdida entre los reconocimientos o entre los lectores del manuscrito; el material cedido, si se cita la fuente, quizá se diluya en un acto de agradecimiento.

Tal es la vida diaria de la ciencia, y por ende su historia. Pero la inmunología tiene de peculiar un hito que supuso un giro copernicano: la teoría de la selección clonal de Frank Macfarlane Burnet, cuyo retrato abre paradigmáticamente el libro. En 1957 la publicaba en un artículo ("A modification of Jerne's theory of anti-body production using the concept of clonal selection") enviado al *Australian Journal of Science*. Su aceptación entre la comunidad no fue un camino de rosas. Por muchas razones. La primera, porque la proponía un oscuro director de un remoto laboratorio australiano, el hoy famoso Instituto Walter y Eliza Hall. Como él comentaría, sin resentida ironía, se trataba de una teoría espléndida, pues podía rechazarse en cualquier momento, de acuerdo con la epistemología poperiana de la falsación. La segunda, y principal, porque estaba muy asentada la hipótesis del modelo "modular" (del "molde", como también se ha llamado en traducción directa el término "template" inglés) de la interacción antígeno-anticuerpo, según la cual era el antígeno quien, a raíz de su presencia en el cuerpo, provocaba la síntesis del anticuerpo a medida.

Gráficamente, el patrón del anticuerpo lo cortaba el antígeno. Burnet hace más plebeyo el sistema y opina que existen muchos trajes de confección en el muestrario y es el antígeno quien se lleva el suyo.

Con una mayor propiedad, aunque la inmunología es una ciencia plagada de metáforas ("respuesta", "reconocimiento", "avidez", "memoria", "rastreo", "tolerancia", "foraneidad", "células asesinas", "células de veto", "células coadyuvantes", "células vírgenes", etcétera), Burnet propone que, cuando un antígeno penetra en el cuerpo, tarde o temprano establecerá contacto con los anticuerpos y "seleccionará" al que más le convenga. (No tiene, pues, nada que ver con la selección darwinista, a pesar del contexto biológico.) Una vez escogida su par, provocará que la célula que lo porta se multiplique y fabrique así el anticuerpo necesario; es decir, se forme un clon.

La relativamente lenta aceptación de la teoría de Burnet va unida al desarrollo de otros frentes de la inmunología; por ejemplo, al avance de la oncología y la genética. El progreso en éstas supuso una mejor descripción de la selección clonal y de la fenómenos inmunológicos, como el de la alergia. Y de eso es de lo que principalmente se habla en esta obra. (L. A.)

Apuntes

Hay proteínas que lo mismo valen para un roto que para un descosido. La estimulación de la insulina provoca que el músculo esquelético que venía degradando el glucógeno pase a sintetizarlo. En ese cambio de rumbo intervienen varias conmutaciones: la quinasa estimulada por insulina se fosforila y activa, en consecuencia, una proteína fosfatasa; ésta a su vez se desfosforila y excita, o activa, la glucógenosintetasa. Se trata, y como tal se ha presentado, de uno de los primeros ejemplos de activación de una fosfatasa por una quinasa, al tiempo que explica el comportamiento paradójico de la insulina. Esta hormona insta la fosforilación de unas proteínas y se opone a la fosforilación de otras.

Con el refinamiento de las técnicas analíticas se ha dibujado el perfil de concentración de titanio en el océano. Nos muestra un gradiente desigual de abundancia de ese elemento, capaz de convertirse así en trazador eficaz de los procesos de transporte químico en mar abierto. El titanio disuelto escasea, se ha observado, en las aguas superficiales y de él se enriquecen las abisales.

Reconocido es el éxito y la validez general del principio de exclusión de Pauli para explicar la regularidad de la tabla periódica de los elementos y los espectros atómicos. A tenor del mismo, sólo puede haber un electrón moviéndose por cada órbita permitida. Desde su formulación en 1925 hasta hace unos pocos años no habían aparecido objeciones de interés, a pesar de que el principio no es absolutamente fundamental para la descripción mecánico-cuántica de la naturaleza. Pero en cuanto se presentó una hipótesis de las partículas que se alejaba levísimamente del principio de exclusión se produjo una avalancha de ensayos experimentales en búsqueda de violaciones del mismo. El último resultado establece una cota superior: la probabilidad de un episodio no pauliano se cifraría en menos de una vez cada 10^{34}.

Los cosmólogos excusan su falta de datos en el uso del condicional: ''si'', ''pudiera ser que...'' (los *ifs* y *maybes* ingleses). Ante la indigencia de razones teóricas, los biólogos echan pronta mano de la explicación adaptativa como panacea universal. Tomemos un ejemplo reciente al azar. Llámase poliploidía la condición de las plantas por la que presentan multiplicado el número haploide de cromosomas (el número de las células germinales). Se ha observado que en determinadas especies suculentas el número multiplicador de cromosomas varía con el tejido y con el estadio de desarrollo del mismo. La justificación ya la hemos avanzado: resistirán así mejor el medio árido donde medran.

Se da por firme la asociación entre puntos calientes de actividad volcánica y plumas térmicas del manto terrestre. Y se citan los ejemplos de Hawai y Yellowstone. Pero, ¿es única esa asociación? ¿También puede atribuirse dicho origen al complejo de puntos calientes de las Azores, en la Dorsal Centro-Atlántica? No debiera. Comienza a hablarse de regiones del manto ricas en agua y de anómala composición química que tienden a fundirse a bajas temperaturas.

Mientras jueces y fiscales de España se quejan de la reciente ley sobre manipulación en la reproducción, tan aventurada cuan imprecisa jurídicamente y ayuna de respaldo científico, el Parlamento alemán ha recabado asesoramiento de embriólogos, genéticos, juristas y moralistas para redactar la suya propia. Recuerda ésta que el destino de los embriones humanos no es otro que el de su implantación uterina, propicia el análisis genético para la intervención en caso de enfermedades y rechaza de plano la libre elección del sexo.

Se admite como dato bien establecido el carácter reductor de la atmósfera de la era arqueana (hace 3800-2500 millones de años). Pero, ¿y si hubiera sido oxidante? A esa hipótesis, ya sin los signos de interrogación, acaba de llegarse tras calcular la productividad primaria global, es decir, la cantidad de carbono orgánico fosilizado en las rocas arqueanas. Parece, en efecto, necesaria, la protoevolución de la respiración aeróbica y de la heterotrofia para mantener los bajos niveles de oxígeno y otras condiciones registradas.

Las marismas de Doñana constituyen uno de los principales enclaves de invernada, reproducción y descanso de aves acuáticas europeas. Pero, ¿cómo pueden abastecer tanta ave, invertebrado y pez que medra en su seno con una producción fitoplanctónica bajísima? Ese milagro de la naturaleza que es el Parque cabe atribuirlo a la biomasa de macrófitos acuáticos, lo mismo sumergidos que flotantes. Ahora bien, tales tapices vegetales tienen su fuerza, y su talón de Aquiles, en la luz que les llega.